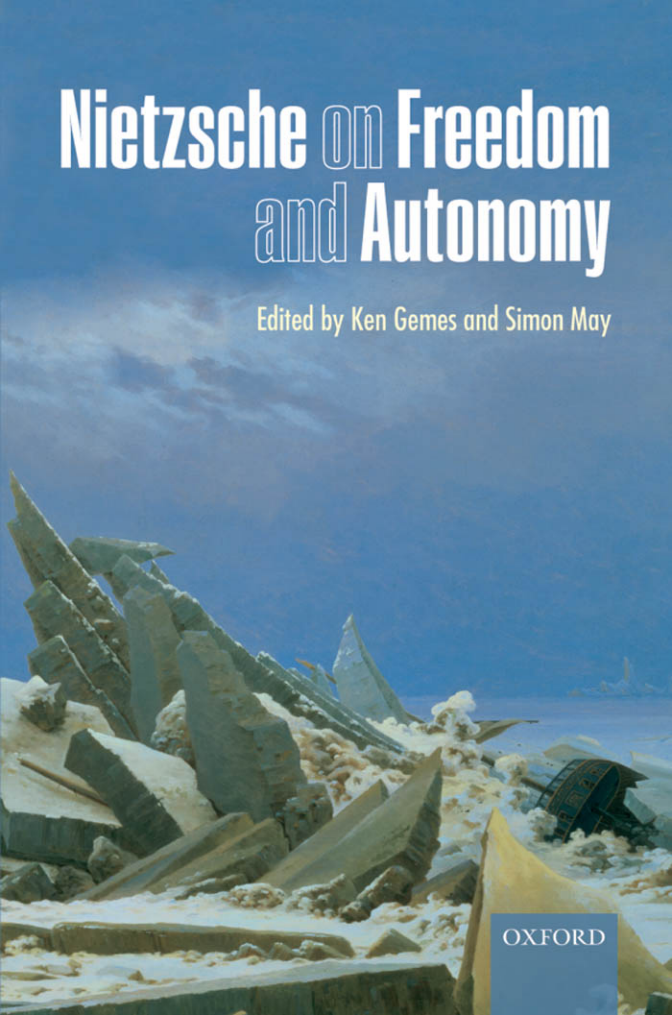
Nietzsche on Freedom and Autonomy

Edited by Ken Gemes and Simon May



OXFORD

NIETZSCHE ON FREEDOM AND AUTONOMY

This page intentionally left blank

Nietzsche on Freedom and Autonomy

Edited by
KEN GEMES AND SIMON MAY

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto
With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam
Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries
Published in the United States
by Oxford University Press Inc., New York
© Ken Gemes and Simon May 2009
The moral rights of the authors have been asserted
Database right Oxford University Press (maker)
First published 2009
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer
British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Nietzsche on freedom and autonomy / edited by Ken Gemes and Simon May.
p. cm.
Includes bibliographical references and index.
ISBN 978–0–19–923156–0
1. Nietzsche, Friedrich Wilhelm, 1844–1900. 2. Liberty. 3. Autonomy (Philosophy)
I. Gemes, Ken. II. May, Simon.
B3318.L43N55 2009
193—dc22
2008054977
Typeset by Laserwords Private Limited, Chennai, India
Printed in Great Britain
on acid-free paper by
Biddles Ltd, King's Lynn, Norfolk
ISBN 978–0–19–923156–0
1 3 5 7 9 10 8 6 4 2

Acknowledgements

We are greatly indebted to Professor Tim Crane and the Institute of Philosophy of the University of London for providing funding and the venue for a conference, *Nietzsche on Self, Agency, and Autonomy*, held on April 7–9, 2006, in which most of the authors here participated and where earlier drafts of their papers were presented. We thank Chris Sykes for invaluable assistance in compiling the index and ensuring consistency in formatting and referencing. Finally we are grateful to Peter Momtchiloff, Philosophy Editor at OUP, for his encouragement, advice, and—not least—patience.

SM, KG

This page intentionally left blank

Contents

Note on Texts, Translations, and References ix
Author Biographies x
Introduction xiii
Simon May

1. Nietzsche, the Self, and the Disunity of Philosophical Reason 1
Sebastian Gardner
2. Nietzsche on Free Will, Autonomy, and the Sovereign Individual 33
Ken Gemes
3. Autonomy, Affect, and the Self in Nietzsche's Project of Genealogy 51
Christopher Janaway
4. How to Overcome Oneself: Nietzsche on Freedom 69
Robert Pippin
5. Nihilism and the Free Self 89
Simon May
6. Nietzsche's Theory of the Will 107
Brian Leiter
7. Nietzsche's Freedoms 127
John Richardson
8. Nietzschean Freedom 151
Peter Poellner
9. Nietzsche's Intentions: What the Sovereign Individual Promises 181
Aaron Ridley
10. Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency 197
David Owen
11. The Eternal Recurrence: A Freudian Look at What Nietzsche Took to Be His Greatest Insight 223
Mathias Risse
12. Nietzsche on the Will: An Analysis of BGE 19 247
Maudemarie Clark and David Dudrick

Index 269

This page intentionally left blank

Note on Texts, Translations, and References

All authors in this volume refer to Nietzsche's works by the same abbreviations, listed below. The particular English translations used by each author are given at the end of his or her chapter, as is the secondary literature to which the author refers.

References to secondary literature are made in the Harvard Style.

A	*The Antichrist*
BGE	*Beyond Good and Evil: Prelude to a Philosophy of the Future*
BT	*The Birth of Tragedy*
CW	*The Case of Wagner*
D	*Daybreak* (also translated as *Dawn*)
EH	*Ecce Homo*
GM	*On the Genealogy of Morals* (also translated as *On the Genealogy of Morality*)
GS	*The Gay Science*
HAH	*Human, All Too Human*
KGB	*Briefwechsel: Kritische Gesamtausgabe*
KGW	*Werke: Kritische Gesamtausgabe*.
KSA	*Sämtliche Werke: Kritische Studienausgabe in 15 Einzelbänden*
NUM	*Unpublished Writings from the Period of 'Unfashionable Investigations'*, trans. Richard T. Gray.
PT	*Philosophy and Truth: Selections from Nietzsche's Notebooks of the Early 1870s*, ed. and trans. Daniel Brezeale.
PTG	*Philosophy in the Tragic Age of the Greeks*, trans. Marianne Cowan.
TI	*Twilight of the Idols: or How to Philosophize with a Hammer*
UM	*Untimely Meditations* (also translated as *Unfashionable Investigations*)
WLN	*Writings from the Late Notebooks*
WP	*The Will to Power*
Z	*Thus Spoke Zarathustra: A Book for Everyone and No One*

Author Biographies

MAUDEMARIE CLARK is George Carleton Jr. Professor of Philosophy at Colgate University. She is author of *Nietzsche on Truth and Philosophy* (1990), co-editor (with Brian Leiter) of Nietzsche's *Daybreak* (1997), and co-translator and co-editor (with Alan Swensen) of Nietzsche's *On the Genealogy of Morality* (1998). She is currently finishing a book with David Dudrick tentatively titled *Nietzsche's Magnificent Tension of the Spirit: An Introduction to Beyond Good and Evil.*

DAVID DUDRICK is Associate Professor of Philosophy at Colgate University. He is the author of 'Foucault, Butler, and the Body', in the *European Journal of Philosophy* (2005), and of 'The Shameful Wisdom', forthcoming *in International Studies in Philosophy*. He is also co-author, with Maudemarie Clark, of a number of papers on Nietzsche. He and Clark are finishing a book tentatively titled *Nietzsche's Magnificent Tension of the Spirit: An Introduction to Beyond Good and Evil*.

SEBASTIAN GARDNER is Professor of Philosophy at University College London. His interests are in post-Kantian philosophy, in particular early German idealism and aesthetics. His publications include *Irrationality and the Philosophy of Psychoanalysis* (1993), *Kant and the* Critique of Pure Reason (1998), and chapters and articles in various books and journals, including the *European Journal of Philosophy*, the *Journal of the History of Philosophy*, and the *Proceedings of the Aristotelian Society*.

KEN GEMES is a Reader in Philosophy at Birkbeck College, University of London, and at Southampton University. His published research is in general philosophy of science and on aspects of Nietzsche's philosophy and appears in various journals including the *Journal of Philosophy*, the *Journal of Philosophical Logic, Philosophy of Science*, the *British Journal for the Philosophy of Science, Noûs, Philosophy and Phenomenological Research*, the *European Journal of Philosophy, Erkentniss*, and *Synthese*.

CHRISTOPHER JANAWAY has been Professor of Philosophy at the University of Southampton since 2005, and before that taught at Birkbeck College, University of London. His research interests fall chiefly in the areas of nineteenth-century German philosophy and aesthetics. His publications include *Beyond Selflessness: Reading Nietzsche's Genealogy* (2007), *Schopenhauer: A Very Short Introduction* (2002), *Images of Excellence: Plato's Critique of the Arts* (1995), and *Self and World in Schopenhauer's Philosophy* (1989); he also edited *The Cambridge Companion to Schopenhauer* (1999) and *Willing and Nothingness: Schopenhauer as Nietzsche's Educator* (1998).

BRIAN LEITER teaches law and philosophy at the University of Chicago, where he is also Director of the Center for Law, Philosophy & Human Values. He is the author of *Nietzsche on Morality* (2002) and *Naturalizing Jurisprudence* (2007) and editor of several books, including *Nietzsche and Morality* (2007) (with Sinhababu) and *The Oxford Handbook of Continental Philosophy* (2007) (with Rosen).

SIMON MAY is College Research Fellow in Philosophy at Birkbeck College, University of London. His research interests are in German philosophy, especially Nietzsche and Heidegger, and in ethics and the emotions. He is author of *Nietzsche's Ethics and His War on 'Morality'* (1999), a collection of his own aphorisms entitled *The Pocket Philosopher: A Handbook of Aphorisms* (1999, second edition 2005, also published in Italian, Spanish and Dutch), *Love: A Philosophical Investigation* (Yale University Press, forthcoming), and other books.

DAVID OWEN is Professor of Social and Political Philosophy at the University of Southampton. He has published nine books, most recently the co-edited collections *Recognition and Power* (2007) and *Multiculturalism and Political Theory* (2007) as well as the authored volume *Nietzsche's Genealogy of Morality* (2007). He is currently working on book projects on the implications of Nietzsche's work for political philosophy and on the ethics of migration.

ROBERT B. PIPPIN is the Evelyn Stefansson Nef Distinguished Service Professor in the John U. Nef Committee on Social Thought, the Department of Philosophy, and the College at the University of Chicago. He is the author of several books on German idealism, including *Kant's Theory of Form* (1982), *Hegel's Idealism: The Satisfactions of Self-Consciousness* (1989), and *Modernism as a Philosophical Problem* (1991). His latest books are *Henry James and Modern Moral Life* (2000), a collection of his recent essays in German, *Die Verwirklichung der Freiheit* (2005), a collection of recent essays, *The Persistence of Subjectivity* (2005), and *Nietzsche, moraliste français* (2005). His research interests include Kant, German Idealism, moral and political theory, contemporary European philosophy, modernity theory, philosophy and literature, and theories of freedom. He is a winner of the Mellon Distinguished Achievement Award in the Humanities, a member of the American Academy of Arts and Sciences, and was recently a fellow at the Wissenschaftskolleg zu Berlin.

PETER POELLNER is Associate Professor of Philosophy at Warwick University. He is the author of *Nietzsche and Metaphysics* (1995) and has published articles on Nietzsche, phenomenology, and non-conceptual content.

JOHN RICHARDSON is Professor of Philosophy at New York University. He is the author of *Existential Epistemology: A Heideggerian Critique of the Cartesian Project* (1986), *Nietzsche's System* (1996), and *Nietzsche's New Darwinism* (2004).

AARON RIDLEY is Professor of Philosophy at the University of Southampton. He works mainly on philosophical aesthetics and on Nietzsche. His publications

include *Nietzsche's Conscience: Six Character Studies from the 'Genealogy'* (1998) and *Nietzsche on Art* (2007). He is currently writing a book on Nietzsche's thoughts about will and action.

MATHIAS RISSE is Associate Professor of Public Policy and Philosophy at the John F. Kennedy School of Government at Harvard University. He works mostly in social and political philosophy and in ethics, but also has a strong interest in nineteenth-century German philosophy, especially Nietzsche. His articles have appeared in journals such as *Ethics, Philosophy and Public Affairs, Noûs*, and the *Journal of Political Philosophy*.

Introduction

Simon May

Why devote a volume to Nietzsche on freedom and autonomy? Nietzsche does not speak often of these topics—as, indeed, he does not speak often of several other topics that he and many of his interpreters take to be central to his philosophy, 'Eternal Recurrence' and 'Will to Power' among them. But freedom and autonomy—the misunderstandings he considers us to have about them and the hopes he has that genuinely 'free spirits' will emerge in a post-Christian future—are deeply important to his philosophical project of revaluing all values. Nietzsche alludes repeatedly to the great difficulty of being free from slave morality and from the cluster of concepts and values that keep it going and that keep us unfree: God, metaphysical absolutes, universal duties, unconditioned values, and the values/affects that distract us from self-creation—above all, perhaps, a certain sort of pity. We moderns might repudiate many of the traditional expressions of slave morality—those that make overt mention of God and metaphysics and universal moral duties—but we still, he holds, cleave to less obvious manifestations of slave morality, manifestations such as substance-predicate dualism, doer-deed dualism, egalitarianism, democracy, universal rights, and the 'atomistic need'—the need to postulate a hard, irreducible subject of actions, objects, and events that endures through time (BGE 12; cf. GM I 13). To cleave to any of this is to be unfree. Unfree to exit slave morality. Unfree to create ourselves.

This is one of the implications of the famous passage on the madman's reaction to the 'death of God' (GS 125). Everyone knows God is dead, says Nietzsche, and people think they are now free of the old order that he personified. The madman's audience, the crowd in the marketplace, ridicules his obsessive focus on this development—his being overwhelmed by it. For them it is already old news, they see themselves as long liberated from all vestiges of superstitious belief. But they think superficially. All they are actually free of is belief in an omnipotent father figure who dispenses laws and judgements, who frightens humans with his power, who is employed as an explanation for why things exist as they do, who promises salvation in return for good behaviour, and who demands feelings of guilt as a reaction to bad deeds. The crowd in the marketplace finds all that laughable, nonsensical, even perverse. But they have not begun to understand two things: what is actually involved in the death of God beyond this superficial loss of traditional faith; and what it would be, instead, to 'become Gods'

ourselves so as 'to be worthy of the deed of killing God'. They think they now follow their own values instead of those dispensed by Scripture and Church, but have they any idea, Nietzsche asks, how to become 'human beings who are new, unique, incomparable, who give themselves laws, who create themselves' (GS 335)? Only a handful of people in modern history have managed this, he repeatedly claims. Goethe, as more than one of our authors points out, is his paradigm of one who is 'strong enough for freedom, a spirit emancipated in the universe' (TI IX 49).

To become free, therefore, the first thing we have to do is to get a sense of how hard and far-ranging the task is. 'How is freedom measured in individuals and peoples? According to the resistance which must be overcome, according to the exertion required . . .' (TI IX 38). If we consider only the power of our contemporary commitment to democratic egalitarianism—according to Nietzsche a paradigmatic product of slave morality—and imagine trying genuinely to discard that commitment, then we can get a sense for the 'resistance which must be overcome' by us modern people to be free in Nietzsche's sense.

Nietzschean freedom involves discovering and digesting the depths to which slave morality goes—the power and pervasiveness of its 'reactive' emotions—and of living, feeling and thinking in a way that 'overcomes' that morality. Those are not separate steps: to digest the realities of slave morality in a lived life—not merely reflectively but in action—is also to overcome it. This is why, for us who are the inheritors of slave morality, whose souls are in many ways structured by slave morality, all overcoming is also a self-overcoming.

SELF-OVERCOMING

If only because slave morality is so ingrained and creating oneself so hard, self-overcoming cannot happen once and for all, like a 'done deal', but is a continuous process. Robert Pippin emphasizes this in his paper. He identifies freedom as a kind of *perpetual* self-overcoming. Freedom, he says, is not a metaphysical capacity to have done otherwise, nor the unconstrained expression of one's identity, but a psychological self-relation—a relation to one's own drives, desires, and commitments. This achievement demands more than merely force of will or self-knowledge or reflective endorsement; it involves a paradoxical form of mindedness, at once affirmative and negative, whole-hearted and ironic (a 'tension of the spirit')—a form of self-relation, Pippin says, that cannot be captured by literary self-creation models.

Promise-making, says Aaron Ridley, is for Nietzsche an arena in which one's self-relation—specifically one's relation to one's drives and commitments—is mercilessly exposed and tested. He argues that there are two distinctive features to the promise-making exhibited by the 'sovereign individual' whom Nietzsche praises in GM II: first, that he refuses to hedge his promises with get-out clauses;

and second, that the most characteristic content of his promises are intentions whose success-conditions cannot be specified independently—cannot, that is, be codified as a set of instructions for successfully executing those intentions. The first feature explains the sovereign individual's 'mastery' over 'circumstances', 'nature', and 'fate'. The second feature explains why the sovereign individual's kind of 'autonomy' makes him 'supra-moral'—why, that is, 'autonomy' and 'morality' are said by Nietzsche to be 'mutually exclusive'.

HIERARCHIES OF DRIVES

Maudemarie Clark and David Dudrick approach this self-relation at the heart of Nietzschean freedom from a different angle. Nietzsche, they hold, sees a person's willing as constituted by the workings (the commanding and obeying) of the normative order of the drives—an order that, rehabilitating a traditional concept in thoroughly naturalistic terms, he calls 'the soul'. In accord with the traditional conception, they suggest, Nietzsche takes the will to be the soul, insofar as the soul is involved in bringing about action. In the course of a detailed exegesis of BGE 19, Clark and Dudrick argue that the self-relation of commanding and obeying—its phenomenology as Nietzsche spells it out in that passage—is what actually constitutes willing for him. Once we recognize this—and that willing not only involves causality but crucially involves the experience of being a causal agent—it becomes implausible, they claim, that BGE 19 aims to show that willing is epiphenomenal, contrary to some other interpretations of this passage, including Brian Leiter's.

RESPONSIBILITY

Central to this self-relation is the reality—and, again, the experience—of being self-responsible. In TI, Nietzsche clearly relates freedom and self-responsibility: 'what is freedom? That one has the will to assume responsibility for oneself . . . .' (TI IX 38). And in the *Genealogy* he asks 'Is there already enough pride, daring, bravery, self-assuredness in existence, enough will of the spirit, will to responsibility, freedom of the will so that henceforth on earth "the philosopher" is truly—possible?'

But, Ken Gemes argues, Nietzsche's concept of responsibility—like the linked concepts of autonomy and freedom—cannot be properly understood if it is tied to the notion of deserts. In order to understand how Nietzsche thinks of responsibility, it is necessary to be clear about how, if at all, he thinks of the will as 'free'. Gemes distinguishes two different notions of free will. Agency free will is intrinsically tied to the question of agency; what constitutes an action as opposed to a mere doing. Deserts free will is intrinsically tied to the question of desert;

of who does and does not merit punishment and reward. Gemes argues that we can render compatible Nietzsche's prima facie conflicting assertions regarding free will—sometimes denying the possibility of free will, while at other times countenancing, even affirming, its existence—by interpreting him as rejecting deserts free will while accepting the possibility of agency free will. Thus Nietzsche advances an original form of compatibilism—one that, in line with other authors in this volume, Gemes sees as taking agency free will to be a rare achievement rather than a natural endowment.

Peter Poellner agrees that Nietzsche's conception of free agency is best construed along compatibilist lines, and he characterizes it as the congruence of a person's effective desires with her values. Valuing, as opposed to bare desiring, is frequently interpreted in the compatibilist literature as a person's *identification* with her effective desires, and Poellner seeks to clarify what the conditions of such identification are. For the mature Nietzsche, he claims, all valuing is ultimately based on affectivity, and some affects, or emotions, are intentional states or, more specifically, perceptions of values in a phenomenological, metaphysically undemanding, sense of 'perception'. According to Poellner, an account of affectivity along these lines enables us to make better sense of a person's identification with some of their desires or motives than projectivist and quasi-realist accounts.

Poellner also holds that, for Nietzsche, a fully autonomous agent is one whose evaluative commitments are independent of purely theoretical metaphysical beliefs, that is, of beliefs the only causally possible justification of which is a priori. Nietzsche's critique of the 'will to truth', for example, involves the thought that even if some such metaphysical beliefs were both justified and incompatible with beliefs implicated in our practical commitments, this should leave our practical self- and world-conception unaffected. The genuinely free or autonomous individual has no significant purely theoretical interests and is consequently indifferent as to whether the life-world which engages her practical concerns is illusory by the lights of a purely theoretical metaphysical inquiry.

Leiter takes a radical stand on the question of what willing means for Nietzsche—and one that wishes to step altogether outside the compatibilist-incompatibilist debate. He analyses Nietzsche's account of the experience we have which leads us (causally) to conceive of ourselves as exercising our will and argues that, according to Nietzsche, the elements of this experience are not causally connected to the resulting action (at least not in a way sufficient to underwrite ascriptions of moral responsibility). Leiter cites recent work in empirical psychology (Libet, Wegner) that, he argues, supports Nietzsche's scepticism that our 'feeling' of will is a reliable guide to the causation of action and his view that willing is determined by a struggle of drives, the outcome of which is a result of 'the underlying type-facts characteristic of that person: namely, which of his various drives happens to be strongest'. Self-mastery is not, therefore, the achievement of a conscious 'self' who contributes anything to

the process, but 'merely an effect of the interplay of certain unconscious drives, drives over which the conscious self exercises no control'.

AFFIRMING THE PAST

Taking responsibility for who we are crucially involves affirming the past—not in the sense of an impotent nostalgia that seeks to escape from the present into an idealized past, but of an affirmation of the causal chain of events that has led to just us. We are, in a sense, our pasts, and so to become who we are, we have to take ownership of our history, of the fate that has led to us and, in important ways, is us.

In this vein, Richardson argues that the different and seemingly inconsistent things that Nietzsche says about freedom can be reconciled by seeing that freedom is, for Nietzsche, a certain skill or capacity that has gone through an extended historical development, and is indeed evolving still. There have been three main phases in this history, each involving a kind of freedom that is layered into the next. So initially freedom is just a capacity in the drives: their capacity to unite or synthesize themselves into a stable 'self'. Human prehistory, however, developed the capacity we might call 'agency'—an ability to restrain the drives in order to remember commitments to linguistic and conscious rules; the exercise of this agency is a second kind of freedom. But this agency has not understood itself, since it has missed the character of its subjection to rules. Genealogy, by giving insight into that subjection, makes feasible a third, de-moralized freedom, which is the kind Nietzsche displays and commends to us.

In my own paper I argue that a central condition for attaining freedom in Nietzsche's sense is to overcome that fundamental form of resentment in the history of Judeo–Christian cultures—resentment against time. Time, so Zarathustra suggests, is the 'dimension' of living that makes us feel most powerless (and thus is the greatest stimulus to resentment). Unlike that better-known object of resentment, the 'masterly type', which can be subdued through 'slave morality', and especially through a certain valuation of pity, equality, and truthfulness, there is *nothing* we can do to change the past. So the hardest test of whether we can affirm our own particular life is not merely to recognize the inevitability of the history that has made us what we are, but actively to 'will' it.

But what is the history in which any contemporary Western self is situated and with which it must reckon if it is to achieve sovereignty? The answer, I suggest, is that our history—the history of Judeo–Christian cultures—*is* the history of nihilism. A sovereign life must therefore be a coming-to-grips with the challenge of overcoming nihilism—that is, of overcoming Western history as the history of the ascetic ideal, as the history of the 'will to nothingness'.

Mathias Risse focuses on one particular way of thinking about time—Nietzsche's idea of the 'Eternal Recurrence'—and notes its strange neglect by many who, like the authors of this volume, see Nietzsche's ethics as, in one form

or another, naturalistic. Among the reasons for this neglect, Risse explores the hypothesis that there is actually a tension between components of Nietzsche's thought stressed by naturalist approaches and the idea of 'Eternal Recurrence', and he employs Freudian concepts to ask how, if at all, one could dissolve such a tension. Specifically, he asks why Freud did not have any appreciation for the Eternal Recurrence, despite offering a similarly naturalistic, but rather more detailed, human psychology to Nietzsche's. Risse argues that there is indeed a tension between endorsing Eternal Recurrence on the one hand and Nietzsche's account of the mind as well as of guilt and *ressentiment* on the other. He suggests four ways of resolving this tension — none of which he considers fully satisfactory.

SELF-LOVE

David Owen approaches the self-relation of the free Nietzschean self in terms of a disposition seldom examined in the secondary literature: self-love. Since self-love, Owen maintains, is clearly understood as exemplary of the 'egoistic' by proponents of modern morality, an analysis of Nietzsche's account of the nature and value of this disposition provides a basis for reflecting on both what he intends by the project of restoring goodwill and value to 'egoistic' actions, and the cogency of this project. Owen begins by sketching the contrast between Aristotle and Kant on self-love and self-respect, before turning to argue that Nietzsche's position can be understood as seeking to overcome the tension between the positions of Aristotle and Kant. He argues, first, that Nietzsche is committed to the value of ethical autonomy, which he takes to ground self-respect (in this respect, he is 'Kantian') and, second, that Nietzsche takes self-love to be necessary to the cultivation of ethical nobility (in this regard he is 'Aristotelian'). The crucial move is then the argument that, for Nietzsche, ethical nobility consists in valuing ethical autonomy and, hence, Owen claims, the tension between self-love (love) and self-respect (law) is dissolved.

THE NIETZSCHEAN 'SELF'

But what is, for Nietzsche, the self that is manifesting 'agency' free will, or loving itself, or perpetually overcoming itself, or making promises, or affirming 'eternal recurrence', or taking responsibility for its actions and its past and Western history as the history of nihilism? Chris Janaway and Sebastian Gardner both suggest that there is a disjunction between, on the one hand, Nietzsche's descriptions of the self as a composite of often competing drives, without a unitary subject of thought or action, which in a well-functioning person is organized into a hierarchy, perhaps under an overall master drive, and, on the other hand, the radically free self that might be an agent of the 'revaluation of all values'. Thus,

Janaway remarks that Nietzsche seeks to dissolve the self into a mere structure of drives and affects, the latter being the sole locus of interpretations; but that the notions of 'controlling' one's affects and 'shifting them in and out', which are requisite for 'perspectival knowing', suggest that Nietzsche must after all accept a relatively unified and autonomous self. Janaway sees a positive conception of free will at work in Nietzsche's 'sovereign individual', which embraces a kind of wholeness of character, acceptance of what constitutes and constrains oneself as an individual, and a creation of one's own values. This is a self that possesses the autonomy to align itself with alternative sets of values as opposed to those it has unthinkingly inherited from the tradition of post-Christian morality. Its 'creative agency' must be assumed, Janaway says, or we could scarcely make sense of Nietzsche's 'proposed critique of moral values and his project of learning to think and feel in healthier ways'.

Gardner argues, in a similar vein, that there is a striking lack of fit between the conception of the self that emerges from Nietzsche's *theoretical* discussion of the self and the conception of the self that is presupposed by his *practical* philosophy. Nietzsche's 'eliminativist-cum-fictionalist' account of the self is inconsistent, Gardner claims, with the substantial and indispensable role that he accords first-person thinking in practical and axiological contexts—a point that Gardner brings out by setting Nietzsche alongside Fichte and Stirner.

The interpretative problem thereby created gives reason to consider the more general question of Nietzsche's view of the relation of the practical and the theoretical, to which the second part of Gardner's paper is devoted. Here he considers in turn two very general, meta-philosophical characterizations of Nietzsche's philosophical project, which construe his view of the practical–theoretical relation in different ways: the naturalistic reading, which models Nietzsche's view of the relation of theoretical and practical thought roughly on that of Hume, and a non-naturalistic, semi-transcendental reading which regards Nietzsche as attempting to construct a unity of philosophical reason. Gardner suggests that, although the latter reading has important advantages over the former, neither is ultimately satisfactory in exegetical respects. In conclusion he proposes a view—a modified version of Habermas's early Nietzsche interpretation—according to which Nietzsche regards philosophical reason as, for us now, incapable of achieving overall unity.

CONCLUSION

A central theme of this book is that Nietzsche embraces notions of free will, autonomy, self-legislation, and responsibility that are wholly naturalistic and that, in doing so, he rejects any notions of free willing that are taken to be causally unconstrained. All the contributors to this volume see Nietzsche as a thoroughgoing naturalist: neither our desires, nor the world they relate to, nor

past and future, nor the ends of ethics, nor anything else about human beings can be described in terms of concepts that abstract from time, space, and the infinite causal relations between this moment, this act, this decision, and all others. There is no room here for noumenal realms where freedom is perfectly unconstrained by these worldly givens, nor for Existentialist moments of choice that are similarly unburdened by the natural causal order. The individual, for Nietzsche, is so inextricably embedded in the natural causal order that strong voluntarism can have no role in any Nietzschean concept of freedom. Given this naturalism, the contributors to our volume hope to shed light on what Nietzsche means when he talks of 'freedom' and 'autonomy' and on what role these concepts play in lives where individuals create and take responsibility for their own values.

1

Nietzsche, the Self, and the Disunity of Philosophical Reason

Sebastian Gardner

> Nous avons psychologisé comme les fous, qui augmentent leur folie en s'efforçant de la comprendre.
>
> Baudelaire, *La Fanfarola*

This paper falls into thee parts. The first section is concerned with Nietzsche's treatment of the self. Here the main point I want to make is that there is a striking lack of fit between the (non-realist or fictionalist) conception of the self that emerges from Nietzsche's theoretical discussion of the self, and the (realist, or at any rate non-fictionalist) conception of the self that is presupposed by his practical philosophy. The problem thereby created gives reason to consider Nietzsche's more general view of the relation of the practical and the theoretical, to which the second section of the paper is devoted. Here again we find, I argue, a disunity in Nietzsche, this time between, on the one hand, a view of the relation of theoretical to practical reason which belongs to modern naturalism, and, on the other, a view of the unity of reason which Nietzsche locates in the Presocratics, but which in modern philosophical terms is recognizably Kantian or post-Kantian. This clash of naturalistic and transcendental conceptions, I propose in the third section, should be regarded as reflecting Nietzsche's view of the limits of (our) philosophical reflection, and in order to clarify and develop this suggestion, I offer a reading of the Third Essay of *The Genealogy of Morals* which, I claim, allows us to understand more exactly how and why Nietzsche's reflection terminates where it does.

The paper therefore works its way through a series of three oppositions which, I argue, are to be found in Nietzsche: (i) fictionalism *vs*. non-fictionalism regarding the self (Section 1), (ii) naturalistic *vs*. Kantian views of the theoretical/practical relation (Section 2), and (iii) naturalistic *vs*. transcendentalist tendencies in Nietzsche's metaphilosophy (Section 3). My claim is that each opposition refers

us on, if we are to make sense of it, to its successor; and that when we come to, and have grasped fully the depth of, the third opposition, there is, in Nietzsche's view, nowhere further for us to go philosophically.

1. NIETZSCHE ON THE SELF

1.1. Nietzsche's Theoretical Conception of the Self

Nietzsche describes the I as an illusion and a fiction. This claim is not qualified in any way—the relevant passages give every indication that it is not just the I of the philosophers, or the I as it may appear in reflection, but the I in *all* respects and contexts of consideration that is held to be illusory. Much of the textual evidence is in the *Nachlaß*,[1] but all of the ideas found there are also

[1] The following (by no means exhaustive) quotations from the *Nachlaß*, given in chronological order, provide a representative selection: 'If *I* have anything of a unity within me [etwas von einer Einheit in mir habe], it certainly doesn't lie in the conscious "I" [. . .] but somewhere else: in the sustaining, appropriating, expelling, watchful prudence of my whole organism, of which my conscious self is but a tool' (WLN p. 2 (n. 34[46])/KGW VII–3.154); 'We *imagine* that what is commanding and highest resides in our consciousness' (WLN p. 6 (n. 34[87])/KGW VII–3.168); 'The concept of the "individual [Individuum]" is false. In isolation, these beings do not exist: the centre of gravity is something changeable; the continual *generation* of cells, etc., produces a continual change in the number of these beings' (WLN p. 8 (n. 34[123])/KGW VII–3.182); 'What separates me most deeply from the metaphysicians is: I don't concede that the "I" is what thinks. Instead, I take the *I itself to be a construction of thinking* [*eine Construktion des Denkens*], of the same rank as "matter", "thing", "substance", "individual", "purpose", "number": in other words to be only a *regulative fiction* [*regulative Fiktion*] with the help of which a kind of constancy and thus "knowability" is inserted into, *invented into*, a world of becoming [. . .] but up to now philosophers have believed, like the "common people", that in "*I* think" there lay something or other of unmediated certainty and that this "I" was the given cause of thinking [. . .] However habituated and indispensable this fiction may now be, that in no way disproves its having been invented: something can be a condition of life and *nevertheless be false*' (WLN pp. 20–1 (n. 35[35])/KGW VII–3.248); 'auch das Erkennen, das absolute und folglich auch das relative, ist ebenfalls nur eine Fiktion! Damit fällt denn auch die Nöthigung weg, ein Etwas das "erkennt", ein Subjekt für das Erkennen anzusetzen [. . .] diese noch von Kant nicht gänzlich aufgegebene Mythologie [. . .] diese Mythologie hat nunmehr ihre Zeit gehabt' (KGW VII–3.342–3); 'Starting point the *body* and physiology: why?—What we gain is the right idea of the nature of our subject-unity [die richtige Vorstellung von der Art unsrer Subjekt-Einheit]—namely as rulers at the head of a commonwealth, not as "souls" or "life forces" [nicht also "Seelen" oder "Lebenskräfte"]' (WLN p. 43 (n. 40[21])/KGW VIII–3.370); 'Seien wir vorsichtiger als Cartesius, welcher in dem Fallstrick der Worte hängen blieb. Cogito ist freilich nur Ein Wort: aber es bedeutet etwas Vielfaches: manches ist vielfach und wir greifen derb darauf los, im guten Glauben, daß es Eins sei' (KGW VIII–3.371); 'The assumption of the *single subject* [des *Einen Subjekts*] is perhaps unnecessary; perhaps it is just as permissible to assume a multiplicity of subjects on whose interaction and struggle our thinking and our consciousness in based? A kind of *aristocracy* of "cells" in which mastery resides? Certainly an aristocracy of equals which together are used to ruling know how to command? *My hypotheses*: The subject as multiplicity [das Subjekt als Vielheit]' (WLN p. 46 (n. 40[42])/KGW VII–3.382); 'the "soul" itself is an *expression* of all the phenomena of consciousness which, however, we *interpret as the cause of all these phenomena* ("self-consciousness" is a fiction [ist fiktiv]!')' (WLN p. 60 (n. 1[58])/KGW VIII–1.21); the I is '*not* the same thing as the unitary government of our being!' and 'is, after all, only a conceptual

presented, with no diminution of force, in published writings,[2] especially the first part of *Beyond Good and Evil*, 'On the prejudices of philosophers'; it can be ruled out, therefore, that the view is merely one that Nietzsche was trying on for size.

In one regard Nietzsche's view that selfhood is illusory follows directly from his general view that consciousness is epiphenomenal: Nietzsche thinks that we take the I of self-consciousness to be a *cause*—we construe its relation to our actions in agent-causationist terms; thus if consciousness is epiphenomenal, so must be the I.[3] But epiphenomena are not fictions, and Nietzsche's thesis is not just that there is an illusory *dimension* to our awareness of the self—that we

synthesis' (WLN p. 61 (n. 1[87])/KGW VIII–1.28); the 'little word "I" ' is 'a mnemonic token, an abbreviating formula [Erinnerungs-Zeichen, eine abkürzende Formel]', not 'an entity [Wesen]': the 'clever feat in the invention' of the subject, of the "I" [Kunststück in der Erfindung des "Subjekts", des "Ichs"]' consists in taking it to be a cause (WLN pp. 96–97 (n. 2[193])/KGW VIII–1.160); 'the word "I", the words "do" and "done to": these may be the horizons of our knowledge [Horizontlinien unsrer Erkenntniß], but they are not "truths" [keine "Wahrheiten"]' (WLN p. 106 (n. 5[3])/KGW VIII–1.187); the I is a constant 'falsifying medium [fälschenden Mediums]' (WLN p. 107 (n. 5[13])/KGW VIII–1.193); ' "Subject": interpreted from the standpoint of ourselves, so that the I is considered subject, cause of all doing, *doer* [. . .] *No* subject "atoms" [*Keine* Subjekt-"Atome"]. The sphere of a subject constantly becoming *larger* or *smaller*—the centre of the system constantly *shifting* [. . .] No "substance", but rather something that as such strives for more strength' (WLN pp. 158–9 (n. 9[98])/KGW VIII–2.55–56); ' "Thinking", as posited by the theorists of knowledge, simply doesn't occur: it is a quite arbitrary fiction [. . .] The "mind", *something that thinks* [. . .] this conception is a derivative, second consequence of the false self-observation that believes in "thinking": here *first* an act is imagined that doesn't occur, "thinking", and secondly a subject-substratum [ein Subjekt-Substrat] is imagined in which every act of thinking, and nothing else, originates; i.e., *both doing and doer are fictions* [*sind fingirt*]' (WLN p. 222 (n. 11[113])/KGW VIII–2.296); 'we have become quite convinced that our concept of "I" guarantees nothing in the way of a real unity [unsere Conception des Ich-Begriffs nichts für eine reale Einheit verbürgt]' (WLN p. 246 (n. 14[79])/KGW VIII–3.50–1).

[2] In *Beyond Good and Evil*, Nietzsche writes that along with the rejection of 'materialistic atomism' must go '*soul atomism* [der *Seelen-Atomistik*]', i.e., 'the soul as something indestructible, eternal, indivisible, as a monad, as an *atomon*' (BGE 12 p. 20/KGW VI–2.21). Nietzsche adds that it is however not necessary to also at the same time get rid of the soul, as do 'clumsy naturalists', for the 'soul-hypothesis' can be refined: 'such conceptions as "mortal soul", and "soul as subjective multiplicity [Seele als Subjekts-Vielheit]", and "soul as social structure of the drives and affects [Seele als Gesellschaftsbau der Triebe und Affekte]" want henceforth to have citizens' rights in science' (BGE 12, pp. 20–1/KGW VI–2.21). The critique of soul here in *Beyond Good and Evil* section 12 is resumed in the analysis of will in BGE 19 in terms of 'a commonwealth' in which 'the governing class identifies itself with the successes of the commonwealth [. . .] a social structure composed of many "souls" [Gesellschaftsbaus vieler "Seelen"]' (BGE 19, p. 26/KGW VI–2.27). (To which corresponds the passage in the Nachlaß, where the 'synthetic concept "I" ' is said to allow us to '*disguise from ourselves*' the duality [Zweiheit] of willing (WLN p. 36 (n. 38[8])/KGW VII–3.334).) See also *Z*, Part One, 'Of the Despisers of the Body': 'You say "I" and you are proud of this word. But greater than this—although you will not believe in it—is your body and its great intelligence, which does not say "I" but performs "I" ' (Z: 62/KGW VI–1.35), and TI: 'The Four Great Errors', 3: 'And as for the ego [das Ich]! It has become a fable, a fiction [Fiktion], a play on words: it has totally ceased to think, to feel and to will!' (TI: 59/KGW VI–3.85).

[3] e.g.: GS 354, pp. 297–300/KGW V–2.272–5; BGE 17, p. 24/KGW VI–2.24–5; TI: 'The Four Great Errors', 3, p. 59/KGW VI–3.85; WLN pp. 2–3 (n. 34[46])/KGW VII–3.154; WLN p. 60 (n. 1[61])/KGW VIII–1.22; WLN p. 60 (n. 1[75])/KGW VIII–1.25; WLN p. 4 (n. 34[54])/KGW VII–3.157.

have false beliefs *about* it—but that its very *existence* is a non-accidental illusion. Nietzsche's fictionalism about the self goes beyond his epiphenomenalism about consciousness, and it forms an interesting contrast with anti-realist views of the self found in the contemporary anglophone philosophical context. In Dennett, for example, the 'I' is regarded as a theoretical posit which is introduced in consequence of taking up the intentional stance, as a kind of conceptual corollary of belief and desire explanation, somewhat in the way that the laws of mechanics warrant the positing of a centre of gravity in physical objects. In Nietzsche by contrast the I is explained in terms of drives or interests, which *lack* at root the cognitive, explanatory character assumed by Dennett. (1) First, our belief in the I is regarded by Nietzsche as a manifestation, perhaps the most fundamental, of our general interest in discovering in the world a platonistic, non-temporal order and structure, an interest which Nietzsche analyses sometimes as a functional matter of filtering the stimulus input to organisms and on other occasions, more deeply (and perhaps in a slightly different respect), as a wishful, *ressentiment*-motivated response to the problem of suffering.[4] (2) Second, the I is explained in terms of power relations between members of the composite of forces which Nietzsche proposes, in contrast with the unifying representation of self-consciousness, as the true structure of psychological reality:[5] Nietzsche supposes (a) that the feeling of I-ness, the sense that the I has of itself as something set over and above the psychological states that it 'owns', is a disguised sensation of power, the subject of which is in reality the presently 'dominant' or commanding member of the psychological composite; and (b) that the 'I' represents to the psychological composite the target of joint action proposed by the commanding unit of will to power, or the norm which it legislates for the collective. (3) Third, and connectedly, the fictional concept of the I has its chief importance, for Nietzsche, in relation to agency, and the distinction which it implies of doer from deed he regards as a case of motivated error, of taking the same event twice over, once as cause (doer, substratum) and once as effect (thing done), all in the service of the slavish fiction of free will.[6]

[4] '*On the psychology of metaphysics* [. . .] These conclusions are *inspired by suffering*: at bottom they are *wishes* that there might be such a world; in the same way, hatred of a world that makes us suffer expresses itself in the imagining of a different world, a *valuable* one' (WLN p. 141 (n. 8[2])/KGW VIII–1.337). Nietzsche accords a role to 'grammar' in perpetuating the belief in the I—see, e.g., GM I 13, p. 28/KGW VI–2.293; BGE 17 and 54/KGW VI–2.24–5 and VI–2.71; WLN p. 21 (n. 35[35])/KGW VII–3.248; and KGW VIII–3.371. But I take it that, even if it enjoys some degree of autonomy, grammar is for Nietzsche not a sufficient source of illusion—we are not innocently misled by it—but rather a contributory factor, especially since he suggests in BGE 20, p. 28/KGW VI–2.29, that the prevalence of specific 'grammatical functions' is not autonomous but lies under 'the spell of *physiological* valuations and racial conditions'.

[5] WLN p. 60 (n. 1[58])/KGW VIII–1.21: '*Man as a multiplicity of "wills to power"* [*eine Vielheit von "Willen zur Macht"*]'. See also, e.g., WLN (n. 14[219])/KGW VIII–3.186.

[6] GM I 13, p. 29/KGW VI–2.294–5: 'The reason the subject (or, as we more colloquially say, *the soul*) has been, until now, the best doctrine on earth, is perhaps because it has facilitated that sublime self-deception whereby the majority of the weak, the dying [. . .]'

The self thus emerges as fictive to a special degree: it shares the unreality of all platonistic-metaphysical conceptions, but has an additional fictivity due to its role in masking the real underlying manifold and in the moralization of humanity's self-conception.

1.2. Kantian Criticism of Nietzsche's Theoretical Conception of the Self

Though the details of this picture are very much Nietzsche's, and bound up with his more general anti-metaphysical theses, in its general outline Nietzsche's theoretical view of the self is not unfamiliar, and it is open to some but not all of the criticisms standardly levelled against so-called Humean or Reductionist theories of the self.

One initial point of importance in clarifying Nietzsche's position: there is a weak sense of illusion in which one may describe an ontological commitment as illusory in so far as it naively fails to grasp that the reality of the item in question is not *sui generis* but actually *consists in* the reality of something else, to which it may be reduced. This is not what Nietzsche means. Although, on occasion, Nietzsche says things that might be taken to suggest a reductionism, whereby the I has reference on each occasion of its use to the dominant member of a psychological composite, his claim is *not* that *what we understand* by the I, the semantic content of I-thoughts, can be given in terms of statements about power relations between psychological states. This *cannot* be his view, because he himself affirms explicitly that the representation I has a primitive, transcendental status (in relation to theoretical knowledge claims), which is incompatible with reductionism: he describes our 'belief in the I as substance' as the assumption 'on which the movement of reason' depends—our attribution of reality to all other things, he says, requires the 'soul-superstition'; this superstition is, he says, the foundation of the oldest realism (WLN 140 (note 7[63])/KGW VIII–1.325), 'the basis on which we make everything *be* or understand it to *be* [nach dem wir Alles *sein* machen oder verstehen]' (WLN 77 (note 2[91])/KGW VIII–1.104), and abandoning it would mean no longer being able to think, 'nicht-mehr-denken-dürfen' (WLN 140 (note 7[63])/KGW VIII–1.325). And he says, in words that could have come from Leibniz or Kant, that we 'have borrowed the concept of unity from our concept of "I"—our oldest article of faith. If we didn't consider ourselves to be unities, we would never have created the concept of "thing" ' (WLN 246 (note 14[79])/KGW VIII–3.50).[7] Since Nietzsche's view is not a reductionism in the classical sense but rather (in terms of our contemporary distinctions) a kind of eliminativism, his position is not vulnerable to the sorts

[7] See also WLN p. 129 (n. 7[1])/KGW VIII–1.258; WLN p. 106 (n. 5[3])/KGW VIII–1.187; WLN p. 77 (n. 2[91])/KGW VIII–1.104; TI 'The Four Great Errors', 3, pp. 48–50/KGW VI–3.84–5; and WLN pp. 158–9 (n. 9[98])/KGW VIII–2.55–6.

of epistemologically-oriented criticisms made by anti-Reductionists, when they argue that the I, the essential indexical, plays an irreducible role in our cognitive performances, which become unaccountable if we suppose the 'I' to refer merely to psychological states and their relations.

Nietzsche's account is exposed, however, to another sort of criticism. There are three ways in which eliminativist positions are standardly supported: (i) by arguments for *incoherence* or lack of conceptual and epistemological integrity in the conceptual item targeted for elimination; (ii) by arguments for its explanatory *redundancy* (in relation to some other, incompatible scheme that does the same job more economically); and (iii) by arguing that the item can be *explained away* (by showing that when we see how our belief in the item arises, and what function the belief serves, we will no longer hold it, or at any rate, no longer be able to understand it realistically). Now there are elements of the first and second patterns of argument in Nietzsche,[8] but it is fair to say that the main burden falls squarely on the third—on Nietzsche's suggestions about the motivation for, and the genesis of the error. The problem arises because any genetic explaining-away of the I, no less than a classical reduction, needs to be expressed in terms that do not presuppose the reality of that which is supposed to be explained away. There is, however, little to show, and little reason to think, that this condition is met in Nietzsche's explanation of how the illusion of the I arises from pre-personal interests and forces of thought. What needs to be explained in the case of the I is not some common-or-garden existential belief, but a very strong, highly distinctive and highly complex reflexive unity, a capacity for constituting identity-in-difference, which goes well beyond the capacity for identifying the aggregates and functional unities which can be discovered in the natural world. So the question arises, how, except in the perspective of an I, of something that takes itself to have unity of the self's sort, can a conception of unity sufficient to account for the fiction of the I be formed? (As it might be put: How can the 'idea' of the I *occur* to a unit of will to power or composite thereof—or to anything *less* than an I?[9]) Nietzsche sometimes talks of parts of the psychological complex as seeking to make themselves masters of the complex, in the way that an individual may take charge of a social group. No doubt parts of an aggregate can generate representations of functional integration, which can then constrain the operation of members of the aggregate in a systematic way,

[8] On the first sort of argument, see BGE 16–17, pp. 23–4/KGW VI–2.23–5, claiming that (i) the notion of 'immediate certainty' constitutively associated with 'I will' and 'I think' is a '*contradictio in adjecto*', and that (ii) the description of consciousness as involving 'I think' is an interpretation that rests on a host of epistemologically arbitrary, unsubstantiated and perhaps unsubstantiatable presuppositions.

[9] Nietzsche merely pushes the problem down a level when he uses reflexive locutions to describe the motivation of units of will to power, as in WLN p. 63 (n. 1[124])/KGW VIII–1.36: perspective arises in an organic being when 'not the being *but the struggle itself wants to preserve itself, wants to grow and to be conscious of itself* [sich nicht ein Wesen, *sondern der Kampf selber erhalten will, wachsen will und sich bewußt sein will*]'.

but this sort of structure is a long way off from reproducing the self's unity. In terms of the social analogy, it is as if one were to attempt to explain how a social group can constitute itself as an object of 'we' thoughts by referring to the mere fact of its brute physical domination by one individual member. In short, relations of power may explain the generation of many kinds of mutual representation among psychological items, but not the specific and distinctive mode of representation which is the 'I'.

Now it should be said that this criticism does not amount to a refutation of Nietzsche's picture of the self. For it can be replied, first, that conceding our epistemic dependence on the irreducible representation 'I' does not commit one to understanding it realistically, and second that the 'I' can be explained away as a fiction generated by the power-relations of interacting elements in the psychological composite, on the grounds that functional requirements can (for all we know) give rise to representations whose content *overshoots* the functional motive that is responsible for creating them—concepts may be generated, in response to functional needs, which go beyond what is needed to fulfil those needs. On this basis, a genetic explaining away of the I fiction can free itself from the obligation to account for the (functionally gratuitous) complexity that the I-representation actually displays.

This reply preserves the coherence of Nietzsche's eliminativism about the self, but it is clear that it also means that the argumentative support Nietzsche is able to offer for it is limited, and so that—at least until the argument is resumed on a broader front—Nietzsche is a long way off from having a *refutation* of Kantian or other realist positions about the self. (Relevant at this point is Nietzsche's implicit acknowledgement that, in order to equip himself with an explanatory base that would be adequate to his naturalistic, eliminativist purposes, a hitherto unimagined complexity will need to be attributed to the body[10]—in other words, Nietzsche is aware that he does not have to hand a strong argument of the second type described above, an argument from greater explanatory economy.)

1.3. The Conflict of Nietzsche's Theoretical Conception of the Self with his Practical Conception

That Nietzsche's theoretical view of the self is not free from problems of a sort that we are familiar with in a contemporary context from assessing hard-naturalistic proposals in cognitive science, evolutionary psychology and suchlike, is a matter that I will leave aside for the moment. The point I want to dwell on now is an internal one, concerning the lack of fit between Nietzsche's theoretical view of the self and the view of the self required for Nietzsche's values.

[10] WLN p. 27 (n. 35[55])/KGW: VIII–1.36: philosophers are, Nietzsche conjectures, now 'learning to exchange for' the idea of the soul the 'even more mysterious' idea of the human body. See also WLN (n. 2[69])/KGW VIII–1.90; and WLN p. 77 (n. 2[91])/KGW VIII–1.104; and Z: Part I, 'Of the Despisers of the Body', pp. 61–3/KGW VI–1.35–7.

Now there are two ways in which one might try to make this point, each involving a different construal of the idea that the self is for Nietzsche the locus of value. (1) First, one might think that there is a conflict of Nietzsche's theoretical picture of the self with his practical philosophy, on the grounds that for Nietzsche the self is the *bearer* of value, that it *is* what is valuable. (2) Alternatively, one might think that Nietzsche's theoretical picture of the self stands in conflict with the *formal conditions* for valuing, more precisely, for any valuing of the specific kind that Nietzsche is engaged in promoting.

Now the first version of the criticism is not obviously successful—on the face of it Nietzsche can consistently hold that the self is in reality nothing but a composite of forces and that this composite, when it has the right kind of organization, is what has value.

The second version of the criticism, however, does find its target (and it also reveals a sense in which the first objection is correct). When we look at Nietzsche's descriptions of what is involved in being the kind of subject that, on his account, realizes value—in the sovereign individual passage in the *Genealogy*,[11] throughout *Zarathustra*,[12] and in numerous other places[13]—we are reminded that the 'I' plays for Nietzsche a fundamental, pervasive, and ineliminable role. This occurs in two closely connected but distinguishable ways. (1) First, the subject who values must understand himself—his self—as the *ground* of the value that he affirms. As in the case of a judgement of taste, where consciousness of one's own imaginative and responsive activity is incorporated into the object of the judgement—taste being of course appealed to by Nietzsche on many occasions as a model for valuing—it is part of what it is to entertain and affirm values in the proper non-alienated, explicitly legislative mode, that one's own contribution, the subject's act of *sponsoring*, be understood as constitutive of the 'object' that comprises one's value. (2) Second, there is a reciprocal relation in Nietzsche between valuing, self-creation, and self-affirmation: to determine such and such to be of value is to determine *oneself*, and to affirm oneself *by way of* affirming

[11] GM II 2, pp. 39–40/KGW VI–2.309–10.

[12] 'This is *my* good, this I love, just thus do I like it, only thus do *I* wish the good' (Z: p. 63/KGW VI–1.38); 'Your love of your neighbour is your bad love of yourselves [. . .] The "You" is older than the "I"; the "You" has been consecrated, but not yet the "I": so man crowds towards his neighbour' (Z: p86/KGW VI–1.73); 'Ah my friends! That your Self be in the action [Dass euer Selbst in der Handlung sei], as the mother is in the child: let that be your maxim of virtue!' (Z: 120/KGW VI–1.119); 'O my Will! My essential, *my* necessity, dispeller of need!' (Z: p. 231/KGW VI–1.264); 'Every soul is a world of its own [. . .] For me—how could there be an outside-of-me?' (Z: p. 234/KGW VI–1.268); 'You creators, you Higher Man! One is pregnant only with one's own child./Let nothing impose upon you, nothing persuade you! For who is *your* neighbour? And if you do things "for your neighbour", still you do not create for him!/Unlearn this "for", you creators; your very virtue wants you to have nothing to do with "for" and "for the sake of" and "because". You should stop your ears to these false little words' (Z: p. 301/KGW VI–1.358).

[13] e.g.: 'To be answerable for oneself, and proudly, too, and therefore to have *the right to say "yes"* to oneself' (GM II 3, p. 41/KGW VI–2.310–1); 'My judgement is *my* judgement—no one else is easily entitled to it' (BGE 43, p. 53/KGW VI–2.56).

what one values, and vice versa. (Further points of this sort, regarding the role of the I in Nietzschean practical thinking, can be made with respect to the thought of *one's own life* that the subject contemplating eternal recurrence entertains, and at an even more basic level, concerning one's thought of oneself as having in the very first place a single life to lead.[14])

So there is in Nietzsche's account of practical thought an 'I will'—'I will that this be prized, etc.'—playing a transcendental role roughly analogous to that of the Kantian 'I think' of theoretical judgement. Just as the 'I think' as a *condition* of judgement does not make the I into the *topic* of judgement, so the Nietzschean 'I will' does not mean that all that is ever valued is oneself, as per practical egoism; to the contrary, an element of self-transcendence—a moment of relating to something other, by means of which a relation to self is constituted—is built into Nietzsche's picture. It is a difficult and interesting question what exactly the Nietzschean 'I will' in relation to values comprises: it amounts to more than the practical commitment that is implied trivially in any recognition of a norm, and appears to include an aesthetic or quasi-artistic dimension—the Nietzschean subject will relate to his values in something of the way that an artist relates to his works, as something distinct from himself, yet as an object of pride and ownership, for which the subject claims responsibility or with which it identifies.

In any case, however the Nietzschean 'I will' is to be understood exactly, what is ineliminable from Nietzsche's picture of valuation is the role of the I: not as substratum, nor as equipped with freedom of will in the indeterministic sense, yet as occupying the position of *ground*. Elimination of this I in favour of thinkings ascribed to complexes of will to power or any other units within the psychological composite would produce a profound self-alienation and undermine the normative dimension of valuation, i.e., the possibility of its being thought that the valuation in any sense 'gets things right'.

This allows us to see what is also correct in the first of the two criticisms. While it is true that there is no inconsistency between *our* thinking of some individual as bearing value on account of their psychological structure, if that individual is to think of *himself* as bearing value, then the I-conception is indispensable: Nietzschean man must set value on *himself*, not on some psychological structure. And since Nietzsche's ultimate philosophical purpose lies in forging individuals who set value on (affirm) themselves, it follows that his account of the bearer of value *does* involve attributing reality to the I.

The problem, therefore, is not just that Nietzsche's theoretical view of the self falls short of furnishing the materials with which to construct the self that figures in his account of practical reason, but that it positively contradicts them. If the 'I' is at best nothing but (as he on one occasion puts it) 'a mnemonic token, an abbreviating formula' (WLN 96 (note 2[193])/KGW VIII–1.160), then it

[14] The first-person dimension of Nietzschean valuation is brought out in Bernstein 1991: see esp. 204 and 207.

cannot *also* be what we invoke in our acts of valuation, under a Nietzschean understanding of these. Nietzsche claims that his deconstruction of the self undermines the metaphysics of egoism,[15] but it equally threatens to undermine his view of valuation as reflexive affirmation. To which may be added that, as a matter of historical tendency at least, the theories of practical reason which appear to consort most readily with Nietzsche's theoretical conception of the self lie in the direction of consequentialism, à la Parfit—not to mention the opening that Nietzsche's theoretical no-self-ism appears to give to Schopenhauer.

1.4. The Contrast with Fichte and Stirner

Now I want to underline this point by contrasting Nietzsche with two other figures, namely Fichte and Stirner, both of whom establish a deep interconnection between their respective theoretical claims about the self and the basic claims of their practical philosophy.

Fichte's account of the self (and his philosophical project as a whole) is designed to demonstrate the fundamental unity of theoretical and practical consciousness, of intellect and will, and because Fichte, like Kant, regards morality as having its necessary and sufficient ground in the structure of practical reason, his account demonstrates a necessary connection of the theoretical conception of the self with value. Moral consciousness is treated by Fichte as a privileged source of illumination regarding self-consciousness in general, the reflexive unity of which is understood by him as necessarily connected with moral agency: the autonomy realized in moral agency is a more explicitly articulated, developed version of the very same structure as constitutes self-consciousness in general.

Stirner too upholds a deep and necessary connection between selfhood and practical norms. In Stirner this consists in something much simpler and less argumentatively structured than what is proposed in Fichte. For Stirner there is a simple and direct alignment of a purported transcendental, pre-practical 'fact' concerning the self, with the recommended axiological position. The transcendental fact in question is simply the bare fact that I am myself (that I am not merely an individual, an *Einzelne*, but a unique-and-only one, an *Einzige*), and the practical position which, he claims, follows from this fact is that nothing can figure for me intelligibly as an object of practical concern that is not in some way 'my own', my *Eigentum*, part of my *Eigenheit, das Meinige*.[16]

[15] WLN p. 61 (n. 1[87])/KGW VIII–1.28: the I is 'only a conceptual synthesis—thus there is no acting from "egoism" '.

[16] Regarding the proper course of a human being's development, Stirner writes: 'As I find myself behind things, and that as mind, so I must later find *myself* also behind *thoughts*, namely, as their creator and *owner* [*Schöpfer und Eigner*] [. . .] And now I take the world as what it is to me, as *mine*, as my property [*Eigentum*]; I refer all to myself' (1995, Part I, ch. 1, 'A human life': 17). Regarding the ideal and proper condition: 'Ownness, on the contrary, is my whole being and existence, it is I myself. I am free from what I am *rid* of, owner of what I have in my *power* or what I *control. My*

Fichte and Stirner are pertinent to a consideration of Nietzsche's account of the self in the following ways.

First, their integration of practical and theoretical reason throws into relief the mismatch in Nietzsche between the theoretical account of the self and practical-philosophical objectives, and at the same time Fichte and Stirner draw attention to what this incongruity consists in at bottom—namely, in the fact that there is in Nietzsche's theoretical treatment of the I no trace of a philosophical account of the self that is in the appropriate way *internal*, that expresses how the self is *for itself*. Fichte and Stirner by contrast *do* theorize the self in a way that gives its proper internal image: they attempt to grasp how the self is for itself, to picture it in philosophical theory in the light of its reflexivity. Nietzsche's theoretical picture of the self by contrast is consistently *external*, and can be entertained only for as long as the self is viewed from the outside, whence its discrepancy with the (necessarily internal) practical point of view.

Second, the contrast with Fichte and Stirner raises the question why Nietzsche should choose not to develop a theory of selfhood that is appropriately internal and that would adequately support his value theory. Fichte's and Stirner's values are, of course, not the same as Nietzsche's, but Nietzsche at least has in common with them the methodological idea that self-experience—a radical turning back on and encounter with oneself—is the fulcrum of genuinely authoritative practical thought; and, as the case of Stirner shows, there is no necessary connection between the view that self-consciousness is significant for practical reason, and a moralism of the sort that Nietzsche objects to, since Stirner is not a Kantian (even Nietzsche, as far as I can see, would be hard pushed to detect ascetic or vestigial Christian tendencies in Stirner, whose polemics against Christianity, self-denial, the fetish of autonomy, and so on, are every bit as vehement as Nietzsche's).

A key point of disagreement of Nietzsche with Fichte and Stirner should be noted. On Fichte and Stirner's accounts, practical norms, values, flow from the *pure* I—from the thin transcendental I in abstraction from any content. This is of course emphatically not Nietzsche's view. I will come back to this point in a moment.

1.5. A Unified Nietzschean Theory of the Self

Before proceeding further with the question of why Nietzsche does not aim at the kind of theoretical–practical integration we find in Fichte and Stirner, it is

own I am at all times and under all circumstances, if I know how to have myself and do not throw myself away on others [. . .] have my law in my whole nature, in myself [. . .] *As own* you are *really rid of everything*, and what clings to you *you have accepted*; it is your choice and your pleasure. The *own* man is the *free-born*' (1995, Part II, ch. 1, 'Ownness': 143, 146, 148). Stirner acknowledges the trans-discursive character of the insight into one's ownness, which he says 'is not in any sense an *idea* like freedom, morality, humanity and the like: it is only a description of the—*owner*' (154).

helpful to ask what theoretical conception of the self *would* be appropriate to Nietzsche's practical purposes, and to what extent such a conception could be made consistent with other elements in Nietzsche's philosophical outlook—that is, whether a unified Nietzschean theory of the self is possible in principle.

It is, I think, not hard to see what adjustments to Nietzsche's theoretical discussion of the self are necessary to bring it into line with his practical conception. In the first place, Nietzsche's critique of traditional metaphysical conceptions of the self does not, on the face of it, *need* to proceed any further than Kant's conclusion in the Paralogisms, namely that there is a boundary to be marked and respected between legitimate transcendental claims for the necessity of the I-representation, and illegitimate, transcendentally realistic claims regarding the constitution-in-itself of a corresponding object of this representation.[17] This, on the face of it, would allow Nietzsche the I that he needs for practical purposes, while also leaving space for his explanation of the I-fiction to do work—the fiction explanation can be redirected, not against the I as such, but against the specific, hypostatized, platonized conception of the I that plays such an important role in Christianity and theologically contaminated philosophy. The merely transcendental, un-hypostatized I would retain a degree of (relative) spontaneity, but this is just the spontaneity that Nietzsche needs for his practical outlook, and it falls short of the absolute spontaneity and transcendental freedom that threatens to bring with it the objectionable (to Nietzsche) Kantian package of noumenalist metaphysics. Nietzsche would retain thereby the self needed for his practical purposes while being rid of the self that he objects to as a life-inimical notion.

In the light of this option, Nietzsche appears to have overshot the mark, and to have failed to absorb the Kantian lesson that there is a middle way between 'soul-substratum' and Humean impersonalism; so when Nietzsche sees it to be necessary, in order to deny that there is a 'substratum' with 'being' that constitutes the metaphysically distinct 'doer' of the deed, to affirm that on the contrary 'the doing is everything' (GM I 13, p. 28/KGW VI–2.293), he is going from one false position to its equally false opposite, throwing out the baby with the bathwater.

What is needed in the second place for a unified Nietzschean theory of the self, is something more demanding, and which is not found ready-made in Kant. As noted, while Nietzsche needs to affirm the reality of the I, he differs from

[17] Nietzsche shows in BGE 54, pp. 66–7/KGW VI–2.71, that he is aware of Kant's position, though he seems to understand it in an idiosyncratic way: modern philosophy since Descartes, Nietzsche says, has experimented with alternatives to regarding the 'I' as the condition for the 'think' as the conditioned, and has proposed instead that 'I' is 'only a synthesis which is *made* by thinking [eine Synthese, welche durch das Denken selbst *gemacht* wird]'; so, says Nietzsche, Kant did entertain 'the possibility of a *merely apparent existence* of the subject [die Möglichkeit einer *Scheinexistenz* des Subjekts]' (an idea which Nietzsche then, perhaps in order to impute an ascetic motive, associates with Vedanta).

Fichte and Stirner in denying that the pure I is a sufficient source of practical normativity. What is needed by Nietzsche, therefore, is a conception on which the I is necessarily always, and without prejudice to its reality, i.e., without being merely reducible thereto, the expression of some or other sort of pre-personal configuration: that is, a conception which holds together in a coherent manner *both* the unitary I of self-consciousness *and* the psychological manifold. This would allow it to be maintained *both* that the causality of the I is properly interpreted as involving necessarily the expression of the dominant power-unit in the psychological composite (or whatever functional arrangement Nietzschean theory tells us is involved), *and* that the thought, which the Nietzschean subject must entertain when a power-unit realizes itself successfully, wills values, etc., is an *I*-thought, not the thought *that* such and such a power-unit or whatever prevails presently. Such an account would explain not just why the Nietzschean subject is no more *able* to give up his I than the Freudian analysand can replace first-person thinking with thoughts about an ego and an id, but why it is *right* for him to think in I-terms.[18]

The task that a reconstruction of Nietzsche faces here, is not peculiar to Nietzsche—it appears also in the philosophy of psychoanalysis, as the problem of understanding how the level of commonsense psychology and self-conscious, self-determining subjectivity joins up with that of unconscious motivation.[19] One historical place where one can see this task at least *recognized* and to some degree embarked upon is in Schelling, in his attempt to embed the Fichtean self (the practical I) in a *naturphilosophisch* context (the I as object of theoretical reason and product of nature), and to some extent also in the German Romantics (Novalis, Friedrich Schlegel, Schleiermacher).

So a unified Nietzschean theory of the self is a possibility. But it is not, in fact, what Nietzsche aims to provide. So there is an interpretative puzzle. Various explanations may be suggested for Nietzsche's disunified position. For example, it may be suggested that Nietzsche affirms a radical dissolution of

[18] For these reasons. the discussion in Reginster (2003: 71–4) seems to me not to go far enough. Reginster describes Nietzsche as wanting to construe the self as a 'unity *in* multiplicity' (p. 71) and acknowledges that 'Nietzsche continues to speak of the self as if it were an independent entity, standing over and above the multiplicity of its own drives' (pp. 73–4), but he equates the self's unity-in-multiplicity (weakly, in departure from the German idealist understanding of this notion) merely with the drives' displaying 'a certain degree of coherence and organisation' (p. 71) and affirms that for Nietzsche 'all there is to the self [. . .] are the drives themselves, and their relations with one another' (p. 72). In defence of this apparently Reductionist interpretation, Reginster suggests that talk of the self as an independent entity is 'just a way of speaking' (p. 74), and that Nietzsche 'believes it unnecessary to explain the unity of the self' (p. 72). It is hard to see on what basis Nietzsche, if he allows the self's unity to be something real, can rightfully deny that it requires explanation, and the claim is contradicted by Nietzsche's many attempts, discussed above, to explain away the putative unity of the self. Reginster's discussion proceeds purely from the third-person psychological-explanatory perspective, and does not factor in the requirements of the first-person practical point of view.

[19] The issue is well stated in Tugendhat 1986: 131–2.

the self solely for *therapeutic* ends—because, he supposes, thinking of oneself *as if* one were an aggregate of forces is somehow emancipating (analogous to the way Lacanian psychoanalysis claims that destroying the illusion of the I's centredness is a condition for therapeutic progress). This would make Nietzsche's own theoretical picture of the self a regulative fiction. Or alternatively it may be suggested that Nietzsche considers that he *has to* deny the existence of the I in *all* senses, on the ground that *any* admission of the I as something over and above the conglomerate of psychological states, is *eo ipso* affirmation of a full-fledged Christian soul—an outlook of which there are indeed some signs in Nietzsche,[20] but which it is hard to see how he could make remotely plausible. Or again, it may be proposed, in a historical perspective, that the explanation lies in the way that the later nineteenth-century philosophical environment had occluded the post-Kantian transcendental tradition and made its philosophical resources unavailable to Nietzsche, a key role being played in this process by Schopenhauer on the one hand, and the naturalistic revision-cum-misunderstanding of Kant which prevailed in the first phase of neo-Kantianism on the other.

There is probably some truth to each of these suggestions, but I suggest that in any case, at the same time, the puzzle should be regarded of deeper and more general significance, on the ground that it repeats itself throughout Nietzsche's philosophy. Again and again we find what appear to be puzzling mismatches between different components of Nietzsche's outlook—between Nietzsche's apparent radical scepticism, and his apparent naturalism; between his scepticism, and the objectivity claims (in the psychological if not the normative sphere) that he appears to need for his project of a revaluation of values; between his anti-essentialism and apparent repudiation of all metaphysics, and his apparent metaphysics of will power (and perhaps also of eternal recurrence); and so on. So there is a pattern, and even if through patient analysis, reconstruction of his claims, discrimination between stages of Nietzsche's development, and so on, one can adjust the parts so that a maximally coherent picture emerges, it seems that the overall *centrifugal* character of Nietzsche's philosophy—the way the bits that compose it, though tethered together, seem to want to fly apart in opposite directions—and the consequent fact that the unity of Nietzsche's philosophy is something that we have to labour to establish, in a way that is not the case for any other modern philosopher: this is surely something that calls for explanation (and to say that it is due to the fact that Nietzsche is not a 'systematic philosopher' is, of course, simply to restate the explanandum).

This brings me to the second part of the paper, where I want to see if any progress can be made towards locating this missing metaphilosophical key,

[20] WLN p. 140 (n. 7[62])/KGW VIII–1.325: 'It seems to me important to get rid of *the* universe, unity, any force, anything unconditional; one could not avoid taking it as the highest agency and naming it God.' Nietzsche may also have been impressed by Frederick Alfred Lange's argument (1950, Bk II, ch. 3: 166 ff) that it is contradictory to suppose something 'absolutely simple' that yet 'has ideas'.

and making better sense of the discrepancy between Nietzsche's theoretical and practical views of the self, by considering how Nietzsche understands the relation between theoretical and practical philosophy.

2. THE PRACTICAL–THEORETICAL RELATION IN NIETZSCHE

There are two prominent candidates for interpreting Nietzsche's general view of the practical–theoretical relation.

2.1. The Naturalistic Model

The first is the naturalistic conception of the practical–theoretical relation, where theoretical reason is regarded as constitutively independent from practical reason, and as committed to philosophical naturalism. Practical thought is regarded on this model as constrained by theoretical reason in so far as its operations must proceed within parameters set by the facts about how things are, which the interests of practical reason play no role in determining. Beyond that, various options present themselves for construing values and practical thought. On one version, practical thinking is reducible to, a special case of, theoretical thinking, while on another, practical reason is left to its own devices by theoretical reason. On all versions, however, practical thinking is itself an *object* of theoretical reason in the sense that, it is held, there are psychological or other naturalistic explanations to be given of our acts of valuing, and these explanations are not incidental to our understanding of value but rather tell us what our evaluations, in some good philosophical sense, really amount to. Thus, to take one example, on Brian Leiter's naturalistic interpretation of Nietzsche, which interprets Nietzsche's view of the theoretical–practical relation on the model of Hume, valuation is characterized as an activity of creative 'legislation' which has no rational connection with cognitive activity, in the sense that valuing is not an instance of cognizing and its content not rationally constrained by the results of cognition. Valuing is at the same time, on Leiter's account, something that Nietzschean psychology can explain according to the methods of the natural sciences, as the empirically determined outcome of prior, non-normative efficient causal conditions.[21]

In Section 3 I will say something about the larger topic of Nietzsche and naturalism. The point to be made here is that the naturalistic view leads to a problem concerning the self-understanding of the Nietzschean subject. We need to ask what, on an interpretation such as Leiter's, the Nietzschean

[21] See Leiter 1998: 223–4 and 2002: 95–6.

subject *understands itself* to be *doing*, when it legislates. On Leiter's Humean interpretation, the legislation of the Nietzschean subject proceeds in an axiological limbo, outside any framework of rational norms, while the subject must at the same time understand its legislation as proceeding in accordance with non-normative, natural law. What follows, on the face of it, is that the Nietzschean subject lacks any rational warrant for regarding his valuation as anything more than the expression of a natural force;[22] and this notion—that one could take one's values, not just when viewing oneself sideways on but also *in the very act of* legislating and endorsing them, to be nothing more than the causal effects of pre-normative psychological forces—encounters a problem of sheer inconceivability for subjects in whom the taste for justification is well established. This picture seems to require for its intelligibility a return to the pre-reflective normative innocence of the 'masters' of *Genealogy* I, leaving Nietzsche wide open to the charge of simple flight from the problems of philosophical modernity.

This criticism can be developed by returning to the case of the self and attending to an important difference between Hume and Nietzsche. Nietzsche's dual account of the self has a striking analogy with that of Hume, who famously denies in Book I of the *Treatise* that there is any idea of the self, while proceeding in Books II and III to make essential use of a fairly rich idea of the self in his analysis of the passions and of the mechanism of sympathy that provides the first plank in his moral psychology. Hume preserves philosophical consistency by holding that the idea of the self at work in the passions is a product of *imagination*, thereby imputing illusion to natural consciousness. Now the Humean-naturalistic interpreter of Nietzsche may accordingly suggest that the same solution is available for Nietzsche: it is philosophically consistent to assert both the non-existence of a self for theoretical reason and the necessity of the *assumption* of a self for *practical* purposes.

Whether this account makes Nietzsche's position consistent depends on whether it mitigates the reflexive unintelligibility that, I suggested, afflicts the Nietzschean subject of legislation on the naturalistic construal, i.e. whether it releases the Nietzschean subject from having to simultaneously affirm and deny the reality of his self. Now we can point to a deep reason why carrying over the Humean model to Nietzsche does not and should not be expected to produce the desired result. Disunity of reason in the natural consciousness of the Humean subject never surfaces *for the subject itself*—it comes to light only in Hume's study, and Hume's reflections in his study are forgotten as soon as he steps outside; so the Hume who knows that he has no self and that his reason is disunified, never gets to meet the Hume who thinks he has a self and whose reason is disunited. The results of Humean reflection never issue in a *problem*—no

[22] Leiter also holds, with consistency, that '*qua* conscious self or "agent", the person takes no active part in the process' of self-creation (1998: 255).

moment of doxastic choice ever presents itself. What keeps the situation stable for Hume is the fact that the Humean subject is a conceptually shallow creature, the course of whose practical and evaluative life is determined pre-reflectively by the causality of affect, of a kind which carries a very light theoretical load. And this is exactly *not* the situation of the enlightened Nietzschean subject, who does not receive values passively from nature but is bound to *innovate* them, and who is consequently exposed to the full battery of reflective questions which, in the Humean subject, the operation of the passions obviates the need and leaves no scope for. The wall separating life from the study has come down for the Nietzschean subject, who is too far from nature to be protected by nature against experiencing the disunity of his reason.

It is in addition of high importance in this context that Nietzsche's values are not hedonistic but consistently *anti*-hedonistic. If it were Nietzsche's view that pleasure or desire-satisfaction provides a sufficient ground for value, then arguably there would be no difficulty, or much less difficulty, in seeing how a naturalistic theoretical self-understanding could determine a practical orientation, in so far as the transition from fact to value could be regarded (so the naturalist might argue) as effected, non-deductively yet with a kind of oblique rationality, simply by way of (the subject's assent to) the causal force of inclination, submission to the lure of pleasure. Nietzsche's view, by contrast, is that—while of course evaluation does have for him the kind of live relation to desire which the Kantian excludes—values of the kind that he advocates, *ideals*, are a great deal more than functions, reflections or derivatives of the hedonic upshot of desire-satisfaction, and cannot be reduced to such.

The overall problem with the naturalistic model, in sum, is that its prioritization of theoretical reason, and the naturalistic constriction of theoretical reason's scope, leave the practical perspective short of the rationality that it needs in order to fulfil the exacting task, which on Nietzsche's un-Humean, non-conservative construal of our axiological predicament, practical reason faces. Arguably there are other, non-Humean, more Aristotelian versions of the naturalistic model that may avoid this implication—richer accounts of Nietzschean nature, that build normativity into it, perhaps centred on the concept of will to power[23]—but they create other problems, by virtue of making Nietzsche's 'naturalism' a very much less familiar kind of thing, and in any case they do not promise to help with the puzzle of Nietzsche's disunified view of the self.

2.2. The 'Presocratic' Unity-Of-Reason Model

The second model regards Nietzsche's philosophy as aiming to construct a unity of theoretical and practical reason in a sense to which philosophical naturalism is indifferent, and which is defined by commitment to a *methodological* principle

[23] See Richardson 1996.

whereby practical considerations have a legitimate role to play at a fundamental level in directing theoretical judgement.

Nietzsche's conception of a unity of the theoretical with the practical is found in his early unpublished writings on the philosophy of the Greeks—in the relatively finished text of 1873, *Die Philosophie im tragischen Zeitalter der Griechen* (*Philosophy in the Tragic Age of the Greeks*), and in the collection of notes from 1875 gathered under the title *Wissenschaft und Weisheit im Kampfe* (*The Struggle between Science and Wisdom*).

In the latter Nietzsche defines 'wisdom', *Weisheit*, in contrast to *Wissenschaft* or 'Socratism' (which, he says, damages ethical life: PT 128/KGW IV–1.174), in terms of three features:

> 1. illogical generalizing and rushing towards final ends, 2. the bearing which these results have upon life [Beziehung dieser Resultate auf das Leben], 3. the unconditioned importance which one ascribes to the soul [unbedingten Wichtigkeit, welche man seiner Seele beilegt]: 'One thing is needful'. (PT 128/*KGW* IV–1.174)

The wisdom of Presocratic philosophy is free, Nietzsche says, from 'the detestable pretension to happiness', and is 'related to *art*' (PT 133/KGW IV–1.179): 'Its solution to the riddle of the universe was frequently inspired by art', inclusive of its 'tragic element', which Nietzsche identifies as 'the attempt to understand the world on the basis of suffering' (PT 135/KGW IV–1.182).[24] Presocratic philosophy begins with Thales' revolt against myth, and eventually, with Empedocles and Democritus, the Greeks found themselves, Nietzsche says, 'well on the way towards *assessing correctly* the irrationality and suffering of human existence': they had uncovered a new set of *possibilities* (PT 131/KGW IV–1.177).[25] These possibilities remained unexplored, however, and were later 'covered up' after Socrates aborted the Greek development by dissociating 'life' and 'knowledge', two drives that thereafter increased in power but, because they were no longer 'under a single yoke', struggled increasingly with one another (PT 143–4/KGW IV–1.191–3).[26] Nietzsche says that his *own* 'general task' is 'to show how life, philosophy, and art have a more profound and congenial relationship to each other', and he adds that he rejects, as did the Greeks, the 'false opposition' of the practical and contemplative lives (PT 134/KGW IV–1.180). This idea of practical-contemplative unity is one component of the comparison

[24] See also PT: p. 149/KGW III–4.144: the birth of tragedy has its 'corroboration in the philosophy of its contemporaries'; and PT: p. 157/KGW III–4.24: 'The world that reveals itself [here: in the "philosophical systems of the early Greeks"], is the same as the one created by tragedy.'

[25] See also PTG: 33/KGW III–2.303 and PTG: 40–1/KGW III–2.307–8.

[26] The Presocratic philosopher 'is contemplative-perceptive like the artist, compassionate like the religious, a seeker of purposes and causalities like the scientist' (PTG: 44/KGW III–2.311).

that Nietzsche draws repeatedly and approvingly of the Presocratic philosopher with the artist: 'the artist stands contemplatively above and at the same time actively within his work' (PTG: 62/KGW III–2.325).[27]

What Nietzsche at this stage calls 'wisdom', it may be claimed, is the standpoint of reflection which aims to unify the practical and theoretical perspectives. The philosopher who possesses or aspires to wisdom is therefore *not just* a legislator and commander, a creator of values occupied only with the practical side of things: he *also* aims to take in theoretical reason and to unify its perspective with the practical perspective. In terms of this early schema of Nietzsche's, philosophical naturalism is philosophy conceived merely as *Wissenschaft*, does not embody the unity of practical and theoretical reason, and is the Socratic antagonist of philosophy as *Weisheit*.

In 1875, therefore, Nietzsche sees himself as returning to the Presocratic task of constructing a unity of theoretical and practical reason, and I suggest that there are substantial interpretative gains to be made by regarding this conception of the task of philosophy—where practical or axiological necessity, the needs of life, have a supreme regulative status for reflection—as preserved in Nietzsche's mature writings.[28] It is evident, furthermore, how this point is connected with the notion that there is a buried transcendental dimension to Nietzsche, as I have already intimated in considering what a unified Nietzschean conception of the self would look like. The unity-of-reason outlook which Nietzsche goes back to the early Greeks to recover is, in the modern philosophical context, properly associated with the transcendental tradition. Quite independently of the historical question of the extent to which Nietzsche may have been influenced by and absorbed motifs from Kant and post-Kantianism,[29] Nietzsche's outlook, in so far as he upholds the Presocratic ideal of philosophy, may be situated, for systemic purposes, in a line of descent from Kant—he mobilizes the idea of the primacy of practical reason and, connectedly, of a non-realistic, transcendental or regulative warrant for norms.

What a transcendental reading of Nietzsche would amount to cannot be explored in detail here, but the following are brief suggestions as to what its core elements should comprise. Nietzsche's naturalistic explanations, on a

[27] See also the *Nachlaß* writings, 'Summer 1872—Early 1873', in NUM: 19[1], 3/KGW III–4.5; 19[12], 5/KGW III–4.7; 19[24], 9/KGW III–4.12; 19[36], 13–14/KGW III–4.16–17; 19[39], 16/KGW III–4.19; 19[41], 17/KGW III–4.20; 19[42], 17/KGW III–4.20–1; 19[45] 18–19/KGW III–4.21–2.

[28] Habermas 1999: 'Nietzsche saw that norms of knowledge are not in principle independent from norms of action: there exists an immanent unity of knowledge and interest' (212).

[29] On the historical relation, see Hill (2003). Commentators who, aside from issues of Kantian historical influence, find various transcendental affinities in Nietzsche include Habermas (discussed below), Deleuze (1983), Gerhard (1988), and Pippin (1991 and 1997).

transcendental account, should be regarded as of a conditional and subordinate sort, dependent on and embedded in an overarching context, which cannot be constructed out of naturalistic materials—a transcendentally conditioned and grounded naturalism. The doctrine of will to power may be viewed as an attempt to construct a picture of the world of the same order as those found, on Nietzsche's view, in the Presocratics (especially Heraclitus)—one which reflects a unity of practical and theoretical reason by virtue of its representing the world in a way that, Nietzsche supposes, is most conducive to the realization of higher purposes, the flourishing of higher types and so forth.[30] Plausibly, other central elements of Nietzsche's philosophy—eternal recurrence, radical scepticism, the special privilege he accords to art—are better explained by supposing that Nietzsche is endeavouring to determine what it is rational to think in the light of the overriding necessity of practical existence, than by attributing to him a bedrock commitment to philosophical naturalism. A transcendental view of Nietzsche's philosophy as a whole need not—any more than, I argued earlier, a transcendental account of the self need do so—undermine Nietzsche's anti-Kantianism; to transcendentalize Nietzsche is not to Kantianize him.

The chief limitation of this interpretative model, it would appear, lies in the difficulty that it encounters in accommodating the strain of aggressive hard naturalism that we also find in Nietzsche, and which is exemplified so clearly by his theoretical treatment of the self: Nietzsche does *not* attempt to construct a unified theory of the self, so here is one place where he does *not* do what the unity-of-reason interpretation would lead one to expect him to do. Now various explanations are possible for this fact—I suggested some earlier—and the disunity of Nietzsche's view of the self is not proof that Nietzsche is not in fact *striving* to forge a unity of reason. What it does seem to mean, however, is that, even if Nietzsche does think that unity of reason is a supreme and necessary philosophical goal, and that practical reason has rights in the theoretical domain, he does *not* think that a unity of reason can *in fact* be achieved, or at any rate, he does not himself see a way to achieve it. Another way of putting this is to say that while Nietzsche may be fully prepared to *subscribe* to the principle of the primacy of practical reason, he does not, it would appear, see a way of *applying* this principle to yield the results that the principle is intended to achieve. And the underlying reason for this, it would appear, is that Nietzsche regards the claims of naturalism as too strong to allow theoretical reason to be bent into the shape that practical reason needs it to assume.

[30] And which finds its post-Kantian transcendental equivalents in the conceptions of Nature—the 'higher physics'—of the Herder- and Goethe-inspired German Romantics, Schelling and *Naturphilosophen* of the early nineteenth century.

3. NATURALISM AND TRANSCENDENTALISM IN NIETZSCHE

3.1. Naturalism, Transcendentalism, and the Problem of Consistency

The upshot of the previous section is, then, that the naturalistic model, though it coheres with Nietzsche's denial of the reality of the I, conflicts with his practical presupposition of the self, and more generally frustrates the ambitions of Nietzsche's practical thought; while the Presocratic-cum-transcendental unity-of-reason model underwrites the rationality of the Nietzschean practical perspective, and has a strong claim to capture the general sense of Nietzsche's project, but seems obliged to recognize that Nietzsche feels the pull of naturalism in a way that forbids the realization of a unity of reason in accordance with the interests of practical reason.

There is a strong temptation, particularly when a historical perspective is applied and it is recalled to what a great extent nineteenth-century German philosophy as a whole is dominated by the two vectors of Kantianism and naturalism, and driven by the attempt to square the idealist legacy with natural science,[31] to regard Nietzsche's philosophy as a compound of naturalistic and transcendental elements, which remain in tension with one another, and which fail to carry through on either the transcendental or the naturalistic side in the way required to achieve philosophical consistency. This would explain why, with respect to the self, Nietzsche appears to be a naturalist with regard to theoretical reason, and a non-naturalist with regard to practical reason.

Views of this kind are found in the literature on Nietzsche. On the pro-transcendental version of his inconsistency, Nietzsche's aversion to the Kantian legacy is responsible for his not having come up with the unified theory of the self (among other things) that his practical project requires. This is, roughly, the estimate of Nietzsche offered by Habermas in *Knowledge and Human Interests*, where Nietzsche is read as proceeding from a transcendental starting point, but as led astray by naturalistic misconceptions, through an overvaluation of the rationality of natural science.[32] Conversely, from a consistently naturalistic angle there will be a story to tell about what impeded Nietzsche's formulation of a properly naturalistic world-view; for example, Nietzsche may be regarded, I suggested earlier, as having failed to renounce his attachment to non-naturalistic, autonomistic values of sovereign individuality and so forth that a rigorously naturalistic world-view is unable to make good sense of, and to reconcile himself

[31] See Schnädelbach 1982. [32] See Habermas 1971, ch. 12, and 1999.

to the broadly utilitarian values which are all that a thoroughgoing naturalism can underpin.

However, before the negative assessment of Nietzsche as stumbling on the path to philosophical consistency—that is, the interpretation of Nietzsche as either (i) an inconsistent transcendentalist, (ii) an inconsistent naturalist, or (iii) simply divided between the two camps[33]—is accepted, we should make sure that no more charitable account of his position is available. It is, after all, not as if the assumptions needed to motivate the reading of Nietzsche as either a transcendentalist or naturalist *manqué* are ones that Nietzsche may be thought to have no objection to: to put matters in Habermas's way is to presuppose the viability of weakening and containing within transcendental bounds the epistemological authority of natural science, just as the naturalistic view of Nietzsche's shortfallings implies a confidence in the capacity of naturalism ultimately to make adequate sense of our axiological situation. The third interpretation, for its part, imputes to Nietzsche a jarring and improbable lack of metaphilosophical self-consciousness.[34]

An alternative, fourth view, which offers more to Nietzsche, is to regard his philosophy, not as unwittingly inconsistent and compromised, but instead as *diagnosing* the disunity in philosophical reason, *identifying* it as marking our philosophical horizon, and *displaying* it for the benefit of our self-understanding. On such a view, Nietzsche is consciously attempting to formulate the non-optimal yet *best available* combination of theoretical and practical claims within the reach of our reflection, *given* the irresolvable contradiction which he sees between the demands of theoretical and practical thought—he thus offers *local* unities of reason, in place of the global unity of reason which Nietzsche may

[33] Green (2002) has argued for a view of Nietzsche as failing in his epistemology to harmonize naturalistic and transcendental factors. Much of the current discussion of naturalism and transcendentalism in Nietzsche revolves around his views of truth and knowledge. While it is certainly true that the question whether Nietzsche is a positivist, or (internal) realist, or transcendental idealist, bears closely on the question whether Nietzsche should be classified as a naturalist or transcendentalist, to think that epistemology provides the key or the primary context for deciding the latter question is, in my view, to mistake the centre of gravity of Nietzsche's philosophy (see Gemes 2001). In any case, if only for reasons of space, I have in this paper largely steered clear of these debates. For what it is worth, my sense is that Green is correct, at least in the sense that the same interpretative indeterminacy that I have been discussing re-emerges in Nietzsche's treatment of epistemological topics. In this connection, see also Han-Pile (forthcoming).

[34] One further interpretative possibility, which to my mind lacks plausibility but which for the sake of comprehensiveness should be mentioned, is that Nietzsche provides, or is aiming to provide, some sort of *Aufhebung* of the opposition of naturalism and transcendentalism (this being, in the eyes of many contemporary philosophers, an independently appealing philosophical target). My reading in the next section of *Genealogy* Essay III, if correct, rules this out, but clearly, the task for anyone who wants to make this strong claim for Nietzsche is first to show exactly how Nietzsche achieves, or attempts to achieve, this metaphilosophical synthesis, and second to explain away the relevant apparent disunities in his thought concerning the self and other topics that an *Aufhebung* would have been expected to overcome.

be regarded as having considered it impossible for enlightened late moderns to achieve.

3.2. The Genealogy of Morals, Third Essay

Next I want to try to support and develop the idea that Nietzsche has philosophical insight into the naturalism/transcendentalism antinomy. My claim is that he, in at least at one place in his mature writings, makes it an *object* of philosophical understanding, in a way that, Nietzsche supposes, allows us to recognize the opposition of naturalistic and non-naturalistic self-conceptions as a (presently) untranscendable limit of reflection.

The *Genealogy* is widely supposed to provide clear support for the naturalistic interpretation of Nietzsche.[35] I want to suggest that, to the contrary, we find Nietzsche at the final point of his enquiry, in the concluding sections of the Third Essay, shifting out of the perspective of naturalism.

In section 23 of the Third Essay Nietzsche asks where we can find a will which opposes the ascetic ideal, and notes that it will be said that our modern science supplies this opposing will, a view which Nietzsche rejects as not speaking from 'the abyss of scientific conscience', i.e., not a scientific interpretation of science (GM III 23, p. 116/KGW VI–2.414). This theme is resumed in GM III 25, where Nietzsche expands on the idea that science does not make itself a value, and so does not supply 'the natural antagonist to the ascetic ideal'. On the contrary, Nietzsche maintains, science 'first needs a value-ideal, a value-creating power' given to it from outside, and this is none other than the ascetic ideal (GM III 25, p. 120/KGW VI–2.420). The notion, associated with the naive view of science as counter-ascetic, that science has extinguished the need to which the ascetic ideal answers—that the effect of man's internalization of science could be to render him '*less in need* of a transcendental solution to the riddle of his existence [*weniger bedürftig* nach einer Jenseitigkeits-Lösung seines Räthsels von Dasein]' (GM III 25, p. 122/KGW VI–2.422)—is flatly rejected by Nietzsche. While it is true that man has been degraded, by no means does this signal a release from the transcendental riddle: rather this degradation answers to the same motivation, which gave birth to the ascetic ideal. Since Copernicus, man 'has become animal, literally, unqualifiedly and unreservedly an animal, man who in his earlier faiths was almost God', and his existence has come to look 'arbitrary, idle, and dispensable in the *visible* order of things', but this experience that man now has of himself—as 'on a downward path', 'rolling faster and faster away from the centre', 'into the *piercing* sensation of his nothingness'—is simply a further ascetic experience, which has the following deep rationale: 'Suppose that everything man ''knows'' does not satisfy his desires but instead contradicts them [. . .] what a divine excuse it is to be permitted to lay the guilt for this at the

[35] See, in addition to Leiter, Williams (1993), and Geuss (1997).

door of "knowing" rather than "wishing"!' Man's self-negation through scientific understanding would amount to a '*triumph* for the ascetic ideal' (GM III 25, pp. 122–3/KGW VI–2.422–3).

From this interpretation of the ascendant post-Copernican view that man is 'unreservedly an animal' as serving the ascetic ideal, it does not follow, of course, that philosophical naturalism is, on Nietzsche's view, to be rejected.[36] What it does mean, however, is that naturalism has been reinterpreted at this point by Nietzsche from the standpoint of practical reason, in terms of which it has been evaluated: the perspective of theoretical reason, from which we are able to grasp that our belief in naturalism has itself a naturalistic explanation and justification, has been supplemented by that of practical reason, from which our belief in naturalism appears problematic. However, since it is through naturalism that we have attained the point we are now at, and in any case no alternatives are on the table, the result is only a provisional alienation, a self-distancing, from naturalism, our commitment to which is not cancelled yet has been shown, from the standpoint of value, to aggravate rather than dispel our fundamental problem.

That with the introduction of this new perspective on naturalism in section 25 an important rupture has occurred becomes clear if we recall the starting point of the book. In the Preface, Nietzsche makes it plain that the practical-axiological problem of morality has absolute priority in his investigation. The expectation which the Preface sets up is, accordingly, that naturalism should contribute to the solution of this practical problem. And up until the present point, section 25, naturalism appeared to have been doing this, its theoretical explanations having—so we have been assuming—a practical point: Nietzsche allowed us to read him as applying knowledge of nature in the manner of a physician making a diagnosis, in the tacit expectation that the same body of knowledge can be employed to supply the cure. By the end of section 25, Nietzsche has indicated that this expectation is disappointed. At any rate, no way of utilizing the theoretical self-knowledge that we have gained in the *Genealogy* to fix our axiological situation is now in view.

In section 28, the short concluding section of the Third Essay, the question which was raised first in section 13, namely, how does the ascetic ideal preserve life? is finally given its answer, and it is to be noted that the terms of consideration are here highly simplified. Nietzsche no longer talks about man in terms of historical epochs, or about particular sub-groups or types of man, but about man in general, whom he considers solely in relation to and in terms of the ascetic ideal, such that man appears to be distinguished solely by it, as if it were his essence: he speaks of the meaninglessness of man's suffering, 'the problem of what he meant', as '*the* curse', in the singular, and of the ascetic ideal, as its (hitherto) single solution (GM III 28, p. 127/KGW VI–2.429).

[36] See Leiter 2002: 281–3.

The theme of suffering and its purpose appears earlier in the *Genealogy*, in section 7 of the Second Essay. There Nietzsche advanced the interpretation of the Greek gods as devices created in order to give suffering a meaning and life a justification, along with the idea that what most 'arouses indignation' over suffering is its senselessness. But back in the Second Essay there was only a faint suggestion of the new idea that appears in section 28. Section 28 presents the need for meaning as transcending the hedonic problem of suffering, and as affording a total perspective on man—man is 'defined', even more basically than by the ascetic ideal, by his need for *Sinn*. And this need to find meaning in our suffering, Nietzsche is now claiming, shows that we cannot reconcile ourselves to nature, and is the reason why we do not and cannot take ourselves to be just nature: if we could represent ourselves as being unreservedly animal, then we would not need our suffering to have meaning. Section 28 is therefore the second non-naturalistic moment in the Third Essay, and it makes the following crucial advance beyond the first: while section 25 merely alienates us from naturalism on the limited axiological ground that it intensifies the ascetic ideal, section 28 characterizes us—or at least, it claims that we *characterize ourselves*—in a way that goes beyond naturalistic understanding. It is crucial for this that when the need for meaning is introduced in section 28, no attempt is made to naturalize it: it is not treated at the physiological level, not historicized or made to seem a product of history, and nothing is said about its origin—rather it is presented as standing outside the foregoing historical story of the *Genealogy*. We are thereby invited to interpret all that has preceded in the light of it, that is, to see the history of man traced in the *Genealogy* as a gradual unfolding of this need, a quasi-teleological process explained at the ground level by the need for meaning. Furthermore, we are invited to read Nietzsche's text itself as motivated by it, that is, to regard Nietzsche's own scientific construction of explanations offered in the *Genealogy* as based on, and deriving its legitimacy from, the need for meaning.[37]

It is also important that this new perspective on man is at the same time, at the present point in the text, a *reflexive* perspective. A comparison with *Gay Science* brings this out. In *Gay Science* Nietzsche writes in the language of a natural historian: 'Gradually, man has become a fantastic animal that has to fulfil one more condition of existence than any other animal: man *has to* believe, to know, from time to time *why* he exists; his race cannot flourish without a periodic trust in life—without faith in *reason in life*' (GS 1, p. 75/KGW V–2.46). The perspective on our need for meaning in *Gay Science* is external—its existence is recorded as a further fact of human natural history. In section 28 of the Third Essay, by contrast, this same fact is viewed and taken up from the *inside*: the need for meaning is presented not as the peculiar trait of a *species*, but rather as a

[37] Leiter, as if half acknowledging this point, calls the need 'existential' (2002: 284), but on his account, 'existential' can mean only 'basic natural'.

practical-axiological problem, which section 28, continuously with the Preface and with section 25, makes us experience as *our own*: we confront it practically and not theoretically, directly and not from a distance.

My contention is therefore that Nietzsche concludes the *Genealogy* with the affirmation that we have a need which points beyond nature and which renders a non-naturalistic self-conception inescapable for us; there is therefore within the *Genealogy*, on the face of it, a transition of its terms of explanation from the naturalism of the First and Second Essays to the concluding recognition of a trans-natural perspective in the Third. Assuming this reading to be correct, the question that must now be considered is what this implies for the interpretation of Nietzsche's metaphilosophy. In particular: Is the naturalistic interpretation of Nietzsche unable to do justice to the points in the Third Essay of the *Genealogy* that I have highlighted?

The naturalistic interpretation is not refuted by the fact that Nietzsche regards scientific understanding as motivated fundamentally by a non-cognitive drive, nor by the fact that it expresses the ascetic ideal, nor again, by its inability to independently create value.[38] The key question is instead what reading it can give of section 28 of the Third Essay.

Note that it will not do for the naturalist to observe simply that all that has been shown, strictly, is that Nietzsche accepts the existence in us of a *representation*, the object of which is, or allows itself to be taken to be, non-natural, namely, *Sinn*, and not that Nietzsche affirms the *reality* of the representation's object. While this is true, the challenge is to make naturalistic sense of the non-naturalistic representation, and this *explanandum* is not disposed of just by being redescribed as the content of a psychological state. Pursuing this task, Leiter claims that the ascetic ideal's solution to the 'curse of meaninglessness' referred to in section 28 is to be understood in terms of the following explanatory elements: first, the claim of the First Essay that suffering produces *ressentiment*; second, the later claim (GM III 15) that relief from *ressentiment* is achieved by blame, which facilitates discharge, the effect of which is anaesthetic; and third, the role of the ascetic priest in redirecting blame by designating a new culprit, namely oneself. In this way the ascetic ideal allows suffering to be 'overcome'. What it is, then, for suffering to 'be meaningful', on the naturalistic account, consists simply in a pattern of psychological causation described in a reductive, mechanistic and hedonistic psychology: it is straightforwardly equivalent to our having the means to achieve hedonic relief from suffering by means of an anaesthesia-inducing discharge of affect directed at a fictional blame-object.[39]

There are two grounds for rejecting this account. First, it inverts the order of explanation implied by section 28. Leiter's naturalistic explanation makes the need for meaning an instrumental *effect* of the *hedonic* practical problem which

[38] All of these considerations are acknowledged fully by Leiter (2002).
[39] Leiter 2002: 256–63. See also Leiter 2003: 93–7.

suffering presents: we need suffering to 'mean' something, i.e. to be taken to have a blameable cause, because and only because we need to find some or other way to end the pain of suffering. But in section 28 Nietzsche presents the need for meaning as the *reason*, in a *non*-hedonic sense, why we *cannot* just live with our suffering, why merely accepting or putting up with suffering, agreeing to suffer, is for us a practical impossibility. Leiter's final *explanans* is the *painfulness* of pain; Nietzsche's is the threat of its *meaninglessness*. These are distinct ways for suffering to be 'unbearable', and Nietzsche's interest is in the non-hedonic, *Sinn*-orientated conditions which determine whether we can or cannot bear our suffering.[40]

Second, the concept of meaningfulness, *Sinn*, as applied to life or existence, is treated generally by Nietzsche in a way that either does not imply, or that positively precludes, its hedonistic psychological reduction. For instance, when in section 7 of the Second Essay of the *Genealogy* Nietzsche introduces the theme of the meaning of suffering, his initial discussion of its hedonic dimension is quite distinctly *overtaken* by the non-hedonic language of life's 'justifying itself'. This concern for justification is credited to the Homeric Greeks, who in Nietzsche's terms elude the ascetic ideal, making it impossible to regard this justificatory concern as a *product* of the ascetic ideal's historical development. Nietzsche's claim in the Second Essay is, furthermore, directly continuous with the *Birth of Tragedy*'s problem of the justification of existence, which is quite clearly not to be understood in terms of hedonistic dynamics, just as tragic affirmation is not presented in that text as a mere hedonistic necessity. Again, in the review of the *Genealogy* given in *Ecce Homo*, it is said that the ascetic ideal has been 'the only ideal' until now, when Zarathustra offers an alternative (EH: 114/KGW VI–3.351); and since Zarathustra and the ascetic ideal address the very same need, and the appeal of the Zarathustran ideal is entirely non-hedonistic, Nietzsche must think that what makes some or other ideal necessary for us, namely the need for meaning, cannot be anything hedonic.[41]

[40] To say this, note, is not to deny that it is *also* true that the need for *Sinn* presupposes, in *some* sense, the fact of suffering. Nietzsche's view of their relation is not straightforward, and consideration of the topic is complicated by the fact that meaninglessness is itself an object or form of suffering (man '*suffered* from the problem of what he meant', GM III 28, p. 127/KGW VI–2.429). My suggestion would be that Nietzsche regards suffering in its primary, hedonic sense as *occasioning* the formation of the need for *Sinn*, indeed as *necessary* for its crystallization, but that he regards the need for *Sinn* as then assuming a life of its own, such that, if *per impossibile* we ceased to suffer in the hedonic-Schopenhauerian sense, we would continue to stand in need of *Sinn*. Suffering thus figures twice in Nietzsche's genealogy, and in two different forms: primordially it is a matter of life's minus score on the hedonic, Schopenhauerian balance sheet; later it assumes the additional and independent form of a consequence of our finding life *sinnlos*.

[41] Compare Freud's peremptory treatment of the 'question of the purpose of human life' in *Civilization and its Discontents* (1961: 75–6). The very question is dismissed by Freud as a further manifestation of 'human presumptuousness' and as intellectually surpassed: 'the idea of life having a purpose stands and falls with the religious system'. Freud accordingly turns to the factual-explanatory, external, third-person question of 'what men themselves show by their behaviour to the

My account of Nietzsche's position in the Third Essay so far may seem to be have been heading towards a transcendentalist conclusion. But there is another question to be faced: *How could* Nietzsche coherently refuse to accept that our need for *Sinn* has a full naturalistic explanation? How could he think that there is anything *more* to what we call our 'need for meaning' than there is to the love of God, i.e., a will directed to a fictional object? How could Nietzsche in any way *reject* the assertion of philosophical naturalism, without a counter-assertion of *super*-naturalism?

On my understanding, however, this is the very point at which Nietzsche regards philosophical reflection as coming to a halt, and the Third Essay of the *Genealogy* is intended to show us how and why it must do so. If Nietzsche were to be a consistent naturalist, then he would have to agree that the need for *Sinn* can be explained as some kind of evolutionary or whatever *Nebenwirkung*, to be resolved back into a naturalized, mechanistic, hedonistic psychology. But—if naturalization of the need for *Sinn* were to have the meaning for Nietzsche that it has for the consistent naturalist—Nietzsche would then have to take Freud's line, that the need for *Sinn* cannot be taken with philosophical seriousness, and his practical philosophy would crumble. Because Nietzsche instead holds fast to the internal, practical perspective from which the question of *Sinn* is genuine and ineluctable, he cannot regard the question of whether the need for *Sinn* is 'naturalizable' as a real question; the possibility that our need for *Sinn* is 'really' nothing but another natural drive or accidental by-product of such is not a real possibility; the need cannot be de-validated through an exercise of theoretical reason. But nor, on the other hand—even though the need for meaning *does have* for him the kind of status that is properly called transcendental—can Nietzsche be interpreted ultimately, comprehensively and overarchingly, as a transcendental philosopher: for he does not think that there is anything to be *made philosophically* of our non-naturalistic self-representation—there is for him nothing that makes that representation 'true', or bestows rationality upon our need for more-than-nature, or that can, in any of the usual philosophical ways, be built upon it by way of a justified theoretical edifice; not even a weaker version of Kant's inference from the incongruity of our representations with nature to our possession of a 'faculty of reason' and corresponding 'rational vocation' can be accepted.[42]

purpose and intention of their lives', i.e. 'the programme of the pleasure principle', in the context of an investigation directed to determining empirically the prospects for reducing human suffering. Freud's quick way is unthinkable for Nietzsche—indeed, it is exactly what he rejects in GM III 23–8. The contrast with Freud highlights Nietzsche's commitment to the integrity of the demand for *Sinn*, and Nietzsche's distance from the hedonism, which forms the axiological mainstream of philosophical naturalism.

[42] The peculiar, hybrid character of Nietzsche's thinking, its complex combination of naturalistic and non-naturalistic perspectives, is shown again in his attitude to the problem of evil. Nietzsche does not—as a consistent naturalist can and should—regard the problem of theodicy as disposed

So we are left by Nietzsche with a disunity of philosophical reason, but not, on the interpretation I am suggesting, with an ascription of philosophical inconsistency to Nietzsche: it is not at all that he simply fails to appreciate the conflict or to decide between naturalism and transcendentalism. What makes the aporetic outcome of Nietzsche's enquiry—his discovery that our rightful search for ('Presocratic', global, theoretical-and-practical) unity of reason cannot succeed—less than fatally crippling for his ultimate philosophical ends is the potential independence, Nietzsche supposes, of the purposes of life from philosophical representations. How Nietzsche supposes this independence can be fostered and sustained—and whether the possibility of the condition he envisages can be made sufficiently convincing for us not to be drawn back into transcendental philosophy—is, of course, another story.[43]

REFERENCES

WRITINGS BY NIETZSCHE

References to writings by Nietzsche use the abbreviations given above in 'Notes on Texts, Translations, and References'. Page references are given first to the English translation listed below, from which quotations are taken, and then (prefixed KGW, and in the form, e.g., 'III–2.3', referring to Abteilung, Band and Seite) to Nietzsche, *Werke: Kritische Gesamtausgabe*, ed. Giorgio Colli and Mazzino Montinari, Berlin: de Gruyter, 1967-.

Beyond Good and Evil: Prelude to a Philosophy of the Future (1886), trans. Walter Kaufmann. New York: Vintage, 1966.

Ecce Homo (1888), trans. R. J. Hollingdale (Harmondsworth: Penguin, 1979).

The Gay Science (1882), trans. Walter Kaufmann (New York: Vintage, 1974).

On the Genealogy of Morals (1887), trans. Carol Diethe, ed. Keith Ansell-Pearson (Cambridge: Cambridge University Press, 1994).

[*Nachlaß* 1872–73] *Unpublished Writings from the Period of 'Unfashionable Investigations'*, trans. Richard T. Gray (Stanford, Calif.: Stanford University Press, 1999).

of in *all* its dimensions by the death of God (see Pippin 2003). Nietzsche thinks that a task corresponding to the theologian's task of furnishing a theodicy remains: it needs to be false that the world is too evil for it to have been created by God; it must be possible to avoid taking the fact of suffering as the objection to God's counter-factual existence which we, as we now are, still cannot help but take it to be. This task requires, Nietzsche thinks, a change of direction of fit: it requires not that the world be shown to be 'really thus' but that we *become otherwise*. (The important, non-naturalistic point here being that this necessity is not naturalistically intelligible.)

[43] I have benefited enormously from detailed comments by Ken Gemes and Brian Leiter on earlier attempts to write about the topic of naturalism and transcendentalism in Nietzsche. I am also grateful for comments on earlier partial drafts of the present paper presented at the University of Warwick, the Radcliffe Institute at Harvard, and the workshop on Nietzsche on Autonomy, Agency, and the Self held in London in 2006, and to the Arts and Humanities Research Council and the Philosophy Department of University College London for research leave that enabled the completion of this paper.

[*Nachlaß* 1872–76] *Philosophy and Truth: Selections from Nietzsche's Notebooks of the Early 1870s*, ed. and trans. Daniel Brezeale (Atlantic Highlands, N. J.: Humanities Press, 1979).

[*Nachlaß* 1885–88] *Writings from the Late Notebooks*, ed. Rüdiger Bittner (Cambridge: Cambridge University Press, 2003).

Philosophy in the Tragic Age of the Greeks (1873), trans. Marianne Cowan (Chicago: Regnery/Gateway, 1962).

Twilight of the Idols: or How to Philosophize with a Hammer (1889) [with *The Anti-Christ*], trans. R. J. Hollingdale (Harmondsworth: Penguin, 1968).

Thus Spoke Zarathustra: A Book for Everyone and No One (1883–85), trans. R. J. Hollingdale (Harmondsworth: Penguin, 1969).

OTHER WORKS CITED

Bernstein, Jay (1991). 'Autonomy and Solitude', in Keith Ansell-Pearson (ed.), *Nietzsche and Modern German Thought* (London: Routledge, 192–215).

Deleuze, Gilles (1983). *Nietzsche and Philosophy*, trans. Hugh Tomlinson (New York: Columbia University Press).

Freud, Sigmund (1961). *Civilization and its Discontents* (1930 [1929]), trans. James Strachey, in *Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. 21 (London: Hogarth Press).

Gemes, Ken (2001). 'Nietzsche's Critique of Truth', in John Richardson and Brian Leiter (eds.), *Nietzsche* (Oxford: Oxford University Press, 40–58).

Gerhard, Volker (1988). *Pathos und Distanz. Studien zur Philosophie Friedrich Nietzsches* (Stuttgart: Reclam).

Guess, Raymond (1997). 'Nietzsche and Morality', *European Journal of Philosophy* 5, 1–20.

Green, Michael Steven (2002). *Nietzsche and the Transcendental Tradition* (Urbana: University of Illinois Press).

Habermas, Jürgen (1971). *Knowledge and Human Interests*, trans. Jeremy Shapiro (Boston: Beacon).

____ (1999), 'On Nietzsche's Theory of Knowledge: Postscript from 1968', in Babette E. Babich and Robert S. Cohen (eds.), *Nietzsche, Theories of Knowledge, and Critical Theory* (Dordrecht: Kluwer, 209–23).

Han-Pile, Béatrice (forthcoming). 'Transcendental Aspects, Naturalistic Elements and Ontological Commitments in Nietzsche's Thought'.

Hill, R. Kevin (2003). *Nietzsche's Critiques: The Kantian Foundations of His Thought* (Oxford: Oxford University Press).

Lange, Frederick Albert (1950). *The History of Materialism: And Criticism of its Present Importance* (2nd edn., 1873–75), trans. Ernest Chester Thomas, 3rd edn., with introduction by Bertrand Russell (London: Routledge & Kegan Paul).

Leiter, Brian (1998). 'The Paradox of Fatalism and Self-Creation in Nietzsche', in Christopher Janaway (ed.), *Willing and Nothingness: Schopenhauer as Nietzsche's Educator* (Oxford: Clarendon, 217–57).

____ (2002). *Nietzsche on Morality* (London: Routledge).

____ (2004). 'The Hermeneutics of Suspicion: Recovering Marx, Nietzsche, and Freud', in Brian Leiter (ed.), *The Future for Philosophy* (Oxford: Clarendon, 74–105).

Pippin, Robert (1991). *Modernism as a Philosophical Problem: On the Dissatisfactions of European High Culture* (Oxford: Blackwell).

____ (1997). 'Truth and Lies in the Early Nietzsche', in *Idealism as Modernism: Hegelian Variations* (Cambridge: Cambridge University Press, 311–29).

____ (2003), 'Love and Death in Nietzsche', in *Religion after Metaphysics* (Cambridge: Cambridge University Press, 7–28).

Reginster, Bernard (2003). 'What is a Free Spirit? Nietzsche on Fanaticism', *Archiv für Geschichte der Philosophie* 85, 51–85.

Richardson, John (1996). *Nietzsche's System* (Oxford: Oxford University Press).

Schnädelbach, Herbert (1982). *Philosophie in Deutschland 1831–1933* (Frankfurt: Suhrkamp).

Stirner, Max (1995). *The Ego and Its Own* (1844–45), ed. David Leopold (Cambridge: Cambridge University Press).

Tugendhat, Ernst (1986). *Self-Consciousness and Self-Determination*, trans. Paul Stern (Cambridge, Mass.: MIT Press).

Williams, Bernard, 'Nietzsche's Minimalist Moral Psychology', *European Journal of Philosophy* 1, 4–14.

This page intentionally left blank

2

Nietzsche on Free Will, Autonomy, and the Sovereign Individual

Ken Gemes

1. INTRODUCTION

Regarding the free will debate it is helpful to distinguish two different approaches. According to the first approach the question of free will is intrinsically tied to the question of desert; of who does and does not merit punishment and reward. For simplicity of reference we might describe the notion of free will relevant to this approach as deserts free will. Writers who focus on deserts free will are typically exercised by such questions as whether determinism precludes free will, thereby undercutting all ascriptions of desert. Where deserts free will is at issue the question of whether having done such and such one could have done otherwise is typically seen as being crucial. The intuition here is that if one could not have done otherwise then one should not be punished/rewarded for what one has done.[1] According to the second approach, the free will debate is intrinsically tied to the question of agency; what constitutes an action as opposed to a mere doing? For simplicity of reference we might describe the notion of free will relevant to this approach as agency free will. Writers who focus on agency free will are typically exercised by questions such as what makes for autonomy. Here the question of whether one could have done otherwise need not loom large.

The approaches can of course be tied together. For instance, one might claim that only a doing that involves free will, where one could have done otherwise, can constitute an action and that one merits rewards and punishments only for one's actions. However the two approaches need not merge. One might, for instance, deny that there is free will in the sense traditionally seen as needed for grounding questions of deserts while at the same time claiming that there is free will in the sense traditionally seen as needed for grounding the notion

[1] Interestingly, while the intuition that one should not be punished for what one could not help doing is fairly strong the intuition that one should not be rewarded for what one could not help doing is typically less strong.

of agency and autonomy. It is the principle burden of this essay to argue that this is exactly Nietzsche's position; Nietzsche rejects deserts free will and affirms agency free will. Nietzsche wants to reject the notion that in doing such and such one might have done otherwise, yet he wants to affirm that genuine agency is possible, if only for a select few. It is this that explains why in some contexts he denies free will and in others he positively invokes free will. The denials are denials of deserts free will and the invocations are invocations of agency free will.

The notion of responsibility has long been tied to the notion of free will. Nearly all who write on free will take it that free will is a condition for responsibility. The not uncommon failure to separate deserts free will from agency free will is to some extent aided and abetted by an ambiguity in the term 'responsibility' (in German 'Veranwortlichkeit').[2] To say that so and so is responsible for such and such can mean that they deserve punishment/reward for it. On the other hand, to say that someone is responsible for such and such can simply mean that it was their doing. The first kind of responsibility is that which is intrinsically linked to the question of deserts, the second kind of responsibility is intrinsically linked to the question of agency. It will be helpful then to separate deserts responsibility, the kind of responsibility which is a precondition of deserved punishment and rewards, and agency responsibility, the kind of responsibility that goes with being the effective agent behind a doing. We shall argue that while Nietzsche rejects deserts responsibility he leaves room for a positive account of agency responsibility.[3]

Now it may immediately be objected that Nietzsche is constantly blaming people (Socrates, Jesus, St Paul, Wagner), and often praising others (Napoleon, Beethoven, Goethe, Wagner, Nietzsche), so it is hard to accept that he is not interested in questions of deserts. It is important here to separate the concept pair of punishment/reward from the concept pair blame/praise. It is natural to run these pairs together, since we tend to think of blame as a kind of punishment and praise as a kind of reward. This is to some extent true for our treatment of our contemporaries, and especially in our treatment of children; typically we reward children by praising them and punish them by blaming or scolding them. In order to separate these pairs of concepts it helps to focus on our attitude towards figures that have long been gone. For instance, those of us who admire Socrates may praise him, yet that praise hardly constitutes a reward—in some sense the dead are beyond rewards, though not beyond honours. Those who, like Nietzsche, take Socrates to be a (largely) baleful influence might blame him, yet not take that blame to be any kind of punishment. Ordinarily punishing and rewarding is something we do to living beings, not those who have already passed away.

[2] Literally, answerability, a term which carries a similar ambiguity. One can be answerable for a result simply in the sense that the result was one's doing; or one can be answerable in the sense that one is deserving of punishment or reward for the doing in question.

[3] The distinction between agency responsibility and deserts responsibility is in some ways similar to the distinction Gary Watson (1996) draws between the aretaic perspective and the moral accountability perspective.

Even in dealing with one's contemporaries one may be interested in praising and blaming them yet not conceive of such praise and blame as a punishment or reward. Nietzsche, in fact, is full of praise and blame for figures who were his contemporaries and figures who were long gone before his time. It is my contention that for Nietzsche such praise and blame is part of his taking them to be responsible for various effects but only in the sense of being agency responsible not in the sense of being deserts responsible. That Nietzsche does blame and praise does not show that he is committed to a positive account of deserts responsibility or deserts free will. In the course of this essay another dimension that will be seen as relevant to Nietzsche's use of praise and blame and his rejection of punishment and reward is the dimension of temporality; Nietzsche rejects the notion of punishment and reward partly because they are in a certain sense directed at the past, a matter of giving what has allegedly been earned; whereas, for Nietzsche, praising and blaming is fundamentally future directed. The objective behind Nietzsche's praising and blaming is to increase or decrease various influences. He blames such figures as Socrates, Jesus, Paul and sometimes Wagner, not because he wants to give them their just deserts but because he wants to lessen their current and future influence. Similarly, he praises Beethoven, Goethe, Nietzsche and sometimes Wagner, because he wants to increase their influence on us and the future, albeit if, at times, only as instructive exemplars.

2. NIETZSCHE'S REJECTION OF DESERTS FREE WILL

Some of Nietzsche's most trenchant dismissals of free will and responsibility occur in works of the middle period such as *Human, All too Human*.

> *The fable of intelligible freedom*. . . . Now one finally discovers that this nature [of man] cannot be responsible, since it is completely a necessary consequence and is assembled from the elements and influences of past and present things; consequently one is not responsible for anything, not for his nature, nor his motives, nor his actions, nor his actions nor for his effects. Thereby one achieves the knowledge that the history of moral sensations is the history of an error, the error of responsibility which rests on the error of freedom of the will. (HAH 39)

The principle target of this dismissal is Schopenhauer and, inter alia, Kant. Kant and Schopenhauer, while recognizing the necessity of all that happens in the phenomenal realm, notoriously allow for a freedom in the noumenal world. Nietzsche by the time of *Human All too Human* has rejected his early Schopenhaurian endorsement of the idea of noumenal existence. The general aim of his rejection of free will and responsibility in HAH is to reject the notions of guilt and deserved punishment and reward:

> For he who is punished does not deserve punishment; he is merely being employed as the means of henceforth deterring others from certain action; likewise, he who is being

rewarded does not deserve his reward: for he could not have acted otherwise than he did. (HAH 105)

In the very next section Nietzsche gives what is probably his strongest endorsement of determinism "if one were all knowing one would be able to calculate every individual action".[4] The emphasis on determinism as an argument against free will and responsibility, and the rejection of the notions of desert in punishments and rewards, clearly indicate that what is at stake here is what we have called deserts free will and deserts responsibility.

In works from his later period, while there is less emphasis on determinism, there is a continued vehement rejection of the notion of free will. In *Twilight of the Idols* he simply refers to "the error of free will" (TI IV 7). In the *Anti-Christ* he says

In Christianity . . . Nothing but imaginary causes ('God', 'soul', 'ego', 'spirit', 'free will'—or 'unfree will'). (A 15)

In these and other passages where he disparages the notion of free will it is clear that what is at stake is the notion of a will autonomous from the causal order, an uncaused cause. It is free will in this "superlative, metaphysical sense" (BGE 21) that Nietzsche rejects. Again context makes clear that it is deserts free will that it is at stake here, the kind of free will seen as necessary to underwrite attributions of deserts responsibility, punishments, and rewards.

3. AGENCY FREE WILL AND THE SOVEREIGN INDIVIDUAL

In the second essay of *The Genealogy of Morality* Nietzsche introduces the striking figure of the sovereign individual, describing him as

. . . . autonomous . . . the man who has his own independent, protracted will . . . this master of a free will (GM II 2—Nietzsche's italics).

This affirmative use of free will is in marked contrast to the dismissals of free will noted above. To get a better handle of what is at stake here it is best to return to the very beginning to the second essay which reads,

To breed an animal with the right to make promises—is this not the paradoxical task that nature has set itself in the case of man? is this not the real problem regarding man? (GM II 1)

The text might easily lead the unwary reader to think this is a task already accomplished, leading the reader into a sense of complacent satisfaction, though in fact the text says only that "this problem has been solved to a high degree".

[4] In this passage, in a clear reference to Schopenhauer, he gives the analogy of a waterfall whose wild movements seem unconstrained but are in fact strictly governed by casual laws.

The sense that Nietzsche is talking of past events is heightened when, having first raised this question of nature's task, he concentrates on the pre-history of man, and man's first acquiring of deep memory—memory burnt in by punishment. The task of acquiring memory is one that has been clearly accomplished; it is something that his audience can proudly lay claim to. However, when a few pages later Nietzsche introduces "the end of this tremendous process" as the "sovereign individual", his audience should at least have a glimmering suspicion of whether they themselves are this proud, noble, sounding individual or the "feeble windbags" he despises. Nietzsche describes the sovereign individual in hyperbolic tones clearly not applicable to ordinary individuals. He describes him as one

> who has his own protracted will and the right to make promises and in him a proud consciousness, quivering in every muscle, of what has at length been achieved and become flesh in him, a consciousness of his own power and freedom . . . [and who] is bound to reserve a kick for the feeble windbags who promise without the right to do so. (GM II 2)

It is typical of Nietzsche's deliberately confusing caginess that it is not at first clear whether the sovereign individual is a creature already achieved or one yet to come. The very terms Nietzsche uses to describe the sovereign individual—"proud", "quivering in every muscle", "aware of his superiority ", "like only to himself", "bound to honour his peers"—clearly hark back to the descriptions of the masters of the first essay. Since his audience are meant to identify themselves as the inheritors of slave morality, it is clear that they cannot be identified with this sovereign individual, who, unlike them, is "autonomous and supermoral", a "lord of the free will". The implicit message to his audience is that you are not sufficiently whole to have the right to make promises; you have no free will, but are merely tossed about willy-nilly by a jumble of competing drives, and, hence, you cannot stand surety for what you promise. You can give no guarantee that the ascendant drive at the time of your making a promise will be effective when the time comes to honour that promise. The type of freedom Nietzsche is invoking here does not involve freedom from the causal order, nor is it bound to questions of deserts. Plainly it is tied to the question of what is it to have genuine agency.

4. NIETZSCHE AS A COMPATIBILIST

One gets a sense of Nietzsche's positive account of free will, and its relation to the tradition, by contrasting it with that of David Hume. Hume, a compatibilist, famously argued that "liberty of spontaneity" (free will) is consistent with the denial of "liberty of indifference" (determinism).[5] On Hume's account, one

[5] Cf. *Treatise of Human Nature*, Bk II, sect. 11. Note, I do not mean to suggest here that Nietzsche is best read as explicitly endorsing both the claim that all events are determined by prior

acts freely where that action stems from one's character. Character for Hume is simply glossed as one's deeper dispositions. Here is not the place to canvass the various problems with Hume's notion of character. What is interesting for us is that Nietzsche may be seen as offering a similar naturalistic account of free will, with the very important difference that he gives a much more robust account of character.[6] To have a character is to have a stable, unified, and integrated, hierarchy of drives. This is a very demanding condition that most humans fail to meet:

> In the present age human beings have in their bodies the heritage of multiple origins, that is opposite and not merely opposite drives and value standards that fight each other and rarely permit each other any rest. Such human beings of late cultures and refracted lights will on the average be weaker human beings. (BGE 200)

In the Nachlass from the period of the *Genealogy* Nietzsche explicitly draws the conclusion that

> one should not at all assume that many humans are "people" . . . the "person" is a relatively isolated fact. (KSA 12: 491, my translation)

The sovereign individual, who has a unified, independent, protracted will counts as having a genuine character, being a person. Modern man, who is at the mercy of a menagerie of competing forces, internal and external, has no such character.[7]

Why after so much denigration of the terminology of free will and autonomy does Nietzsche in *The Genealogy* employ it in a positive fashion? Presumably, as a subtle challenge to his readers. Rather than simply arousing his (libertarian) audiences' resistance with flat denials of free will and autonomy in the transcendent sense, or, in the case of (incombatiblist) naturalists, confirming their flat rejection of free will, Nietzsche uses that terminology in a positive, non-transcendent, manner in describing the sovereign individual. He then seeks to unsettle his audience with the uncanny idea that autonomy and free will are achievements of great

events and the claim that free will is possible. Rather he is a compatibilist in the sense that he does not take determinism to be incompatible with free will. As will become clear below, Nietzsche is best read as one who has implicitly realized that the issues involved in making sense of free will do not have a direct connection with determinism. I do believe Nietzsche is committed to the idea that some rare individuals act freely; as for determinism, I do not think he has any real commitment there. More generally I think such metaphysical views were not within his philosophical provenance—he occasionally flirted and dabbled with such theses but did not give them sufficient reflection necessary for genuine commitment. More typically he uses such metaphysical claims as occasional tools to help dislodge various ideas he takes to be harmful, for instance the moralist's obsessions about guilt and responsibility. Nietzsche is a Kulturkritiker and psychologist, perhaps even a moral philosopher, but a metaphysician he is not, nor did he care to be. The one exception to this claim is perhaps his thought concerning the will to power.

[6] An important difference between Nietzsche and Hume is that Hume's account of free will is an account of both agency and deserts free will, or at least an account of our attributions of agency and deserts free will. Nietzsche's account is only an account of agency free will.

[7] For more on this see Gemes (2001) and ch. 1, sect. 5 of the excellent Richardson (1996).

difficulty, achievements which they themselves have by no means attained. While the thought that free will does not exist is disturbing, how much more so is the thought that free will does exist but one does not oneself possess it! What is at stake here is clearly the notion of agency free will and agency responsibility. It is these that are being attributed to the sovereign individual.

5. AGENCY FREE WILL VERSUS DESERTS FREE WILL

Those who take the key issue concerning free will to be the questions of moral responsibility, deserts, the justification of punishments and rewards, will claim that this notion of agency free will relevant to the notion of the sovereign individual is no genuine notion of free will. For them an action is free only if the agent could have done otherwise.[8] But as we have seen there is another way of approaching the free will debate. This other way does not see the debate directly through the question of moral responsibility, but, rather, approaches it from the question of agency. Where one approach begins with the questions "For what acts is one morally responsible? For what acts can one be punished or rewarded?", the other begins with the arguably profounder question "What is it to act in the first place, what is it to be a self capable of acting?". Those who take the questions of moral responsibility and desert as paramount to the free will question, tend to write as if we already have a notion of self and action more or less firmly in place and are only raising the question of whether such selves are ever to be held morally responsible for their actions. The other approach seeks to problematize the very notions of self and action. This is part of the import of the famous dictum from the *Genealogy* that "the doer is merely a fiction added to the deed" (GM I 13). Nietzsche is here questioning our assumption that for every deed there is an individual agent who does the deed.[9] Now Nietzsche does, of course, want to question our practices of punishing and rewarding, our practices of assigning moral responsibility. This is part of his ongoing battle against the dominant Judeo–Christian worldview in which moral responsibility,

[8] Leiter's (2002) provides a well-argued and trenchant account of Nietzsche as a denier of free will. Leiter resists any compatibilist reading on the grounds that any genuine notion of free will must provide a basis for ascriptions of responsibility capable of grounding our practices of punishing and rewarding. Leiter is really insisting that free will can only be deserts free will and that the compatibilist agency free will we have ascribed to Nietzsche is really a wholly revisionary concept of free will, divorced from the "conventional meaning" of free will. Cf. especially Leiter (2002: 99). This reading rests on a rather restricted reading of the history of philosophy. Putting it crudely, and oversimplifying, while the modern Anglo-Saxon tradition may have a convention of interpreting the question of free will largely through the narrow prism of the question of moral responsibility and deserts, the German tradition, including Kant, Hegel, Fichte, and Schelling, has a convention of interpreting the question of free will as much, if not more, through the prism of the question of agency and autonomy.

[9] For some deeds in which individual humans are involved what is behind the deed is merely the will of the herd.

and, in particular, punishment and guilt are key notions. This aim does indeed account for many of his negative comments about free will. But ultimately the more profound Nietzsche wants to raise the question about what exactly it is to be a genuine self. Indeed, his whole attack on the Judeo–Christian worldview is predicated on his belief that it is fundamentally inimical to the development of genuine selves. It is for this reason that for Nietzsche the problem of agency free will takes precedence over the problem of deserts free will; those who take the deserts free will problem as central typically complacently assume that we have a coherent notion of self and agency already in hand.[10]

It is worth pausing to note that Nietzsche, from a purely philosophical point of view, and like the compatibilist Hume, considers free will, in the metaphysical sense of *causa sui*, to be actually incompatible with ascriptions of moral responsibility. In a rather neglected passage from *The Wanderer and His Shadow* 23, "Have the adherents of the theory of free will the right to punish?", Nietzsche presents the argument that if free will in the metaphysical sense really existed it would in fact absolve its bearers from being candidates for punishment:

> The offender is punished because he employs "free-will", that is to say, because he acted without a reason, where he ought to have acted in accord with reasons. Why did he do this? But it is precisely this question which cannot be asked [according to the proponents of *causa sui* free will]: it was a deed without a 'for that reason', without motive, without origin, something purposeless and non-rational.—*But such a deed too ought*, in accordance with the first condition of punishability laid down above [if he is to be punished he must have acted for reasons], *not to be punished*.

Determinist incompatibilists make the claim that there is moral responsibility if and only if there is free will (in the *causa sui* sense) and then go on to reject any such notion of free will. Typically philosophers have concentrated on one side of this biconditional, namely, that which says that if there is moral responsibility then there is free will. The other side, the conditional that if there is free will then there is moral responsibility is typically neglected except by some shrewd compatiblists who point out that *casua sui* free will could not ground responsibility. The fact that Nietzsche rejects this conditional is grist for interpretations that ascribes to him some form of compatibilism. Of course Nietzsche realizes that since so many make the incompatibilist's error of thinking that free will is both necessary and sufficient condition for responsibility, denying free will is a powerful way of dismantling the notion of responsibility. This just shows again Nietzsche's incredible polemical sensitivity. Nietzsche knows that as a means of arguing against responsibility attacking metaphysical free will is an effective tool, even if

[10] It is important to realize that this criticism is not directed specifically against philosophers but rather against what in the *Genealogy* Nietzsche calls "we moderns" who are unknowingly "strangers to ourselves". For more on this see Gemes 2006. That some philosophers, for instance, Kant, Hegel, Fichte, and Schelling were clearly aware of the problematic nature of autonomy and agency—even granting that Nietzsche rejects their accounts of these notions—is besides the point.

it depends on a philosophical error. This just shows how interested Nietzsche was in combating the notion of responsibility. It is not a basis, first, for assigning to him the incompatibilist notion that free will is a necessary and sufficient condition for responsibility, and second for taking him to actually reject all notions of free will, though it is clear he rejects the *causa sui* notion.

To interpret Nietzsche simply as an opponent of deserts free will because he rejects *causa sui* free will, deserts free will, is to emphasize a purely negative aspect, his hostility to Judaeo–Christian notions of responsibility and punishment. This undoubtedly is an important and oft repeated theme in Nietzsche's work. To interpret Nietzsche as giving a positive, naturalist account of agency free will is to emphasize a positive and original aspect, his notion that under the right conditions genuine agency, a truly great achievement, is possible, albeit only for a special few. Furthermore this interpretation helps us to properly appreciate the famous passage in *Gay Science* 125, where Nietzsche's "madman" tells us that we must ourselves become Gods to be worthy of the deed of killing God. To become Gods is to be autonomous, self legislators who are not subservient to some external authority, be it a God, the *sumum bonum*, or an, allegedly, universal moral law.

6. NEITHER FREE OR UNFREE; BUT STRONG OR WEAK WILLS

In some prominent places where Nietzsche disparages the notion of free will he also takes care to disparage the notion of unfree will, as we saw in the quotation above from *The Antichrist*. In the famous passage from BGE were he rejects free will in the superlative metaphysical sense he concludes

> Suppose someone were thus to see through the boorish simplicity of this celebrated concept of "free will" and put it out of his head altogether, I beg of him to carry his "enlightenment" a step further, and also put out of his head the contrary of this monstrous conception of "free will": I mean "unfree will," which amounts to a misuse of cause and effect. (BGE 21)

Why does Nietzsche do this? Clearly a central reason why Nietzsche rejects the traditional notion of free will is that it typically functions to instil guilt, thereby fostering a passive attitude to the world (the rule of the "Thou shall not"s). But he is also aware that the new modern notion of unfree will, by fostering a fatalistic pessimism, what in BGE, he calls "a fatalism of the weak willed", can also be inimical to active worldly engagement. For Nietzsche "unfree will is mythology, in real life it is only a matter of strong and weak wills"[11] (BGE 21).

[11] Gudren von Tevenar has pointed out that given that talk of strong and weak wills suggests the notion of a sliding strength scale of wills, perhaps a more nuanced version of Nietzsche's account of

It is when a strong will takes command, orders and organizes lesser drives that a genuine self can emerge. This is implicit in his accounts of Wagner's development.

> The dramatic element in Wagner's development is quite unmistakable from the moment when his ruling passion became aware of itself and took his nature in its charge: from that time on there was an end to fumbling, straying, to the proliferation of secondary shoots, and within the most convoluted courses and often daring trajectories assumed by his artistic plans there rules a single inner law, a will by which they can be explained. (UM III 2)[12]

The story of Wagner's achievement of a higher unity borne from some master drive is of course the story Nietzsche would repeat about himself in the dramatic section of *Ecce Homo* where Nietzsche elaborates the subtitle of that work "How One Becomes What One Is":

> To become what one is, one must not have the slightest notion of what one is. . . . The whole surface of consciousness—consciousness is a surface—must be kept clear of all great imperatives. . . . Meanwhile the organizing "idea" that is destined to rule keeps growing deep down—it begins to command; slowly it leads us back from side roads and wrong roads; it prepares single qualities and fitnesses that will one day prove to be indispensable as a means towards the whole—one by one, it trains all subservient capacities before giving any hint of the dominant task, "goal", "aim", or "meaning". (EH, "Why I am So Clever", 9)

Nietzsche, as we shall soon see, often gives similar accounts with respect to Goethe's achievement of self creation.

A helpful picture: according to Nietzsche most humans, being merely members of the herd, are merely passive conduits for various disparate forces already existing and operating around them. Some individuals, due perhaps to conscious design but more likely due to fortuitous circumstances, actively collect, order and intensify some of those disparate forces and create a new direction for them, thereby, in fortuitous circumstances, reorienting, to some degree, the whole field of forces in which we all exist. It is these individuals according to Nietzsche who deserve the honorific person, who by imposing their strong will exercise a form of free will and genuine agency.

In an interesting passage in *Beyond Good and Evil* Nietzsche characterizes the generation of his day as having

autonomy and agency free will would allow that these also come in degrees. The point is well taken, however Nietzsche's strong concern with the question of how to awaken his audience from their complacent assumptions, militated in favour of the more rhetorical effective presentation in terms of agency free will as something either wholly present (for instance, in the sovereign individual) or something wholly absent (for instance, in modern windbags).

[12] Of course Nietzsche would later take a less favourable view of Wagner.

Inherited in its blood diverse standards and values, everything is unrest, disturbance, doubt, attempt, the very virtues do not allow each other to grow and become strong; balance, a centre of gravity, and perpendicular poise are lacking in body and soul. But what becomes sickest and degenerates most in such hybrids is the will: they no longer know independence of decisions and the intrepid sense of pleasure in willing—they doubt the "freedom of the will" even in their dreams. (BGE 208)

The implication here is that strong individuals who have a "centre of gravity" would not doubt that they have freedom of the will. On this picture having free will is not a matter of being free of necessity, but rather acting from a kind of inner necessity stemming from this centre of gravity. It is just this type of necessity that prompted Luther to famously say "Here I stand, I cannot do otherwise".[13]

Nietzsche's positive account of free will and agency finds its echo in his many positive pronouncements about responsibility in *Ecce Homo* and especially in *Beyond Good and Evil*. Typically these positive pronouncements are forward looking; they involve responsibility towards the future. In the *Ecce Homo* passages they evince a near megalomaniacal sense of possible influence on the future. Thus in *Ecce Homo*, "Why I am So Clever", Nietzsche speaks of having a "responsibility for all millennia after me" and in the same work in the section "The Case of Wagner" he says, "an indescribable responsibility lies upon me . . . in that I bear the destiny of mankind on my shoulders". These hyperbolic pronouncements find a more toned down expression in his general pronouncements about responsibility in *Beyond Good and Evil* where he continually emphasizes the duty of the philosopher to take responsibility for the future, for it is "[t]he philosopher as *we* understand him, we free spirits—as the man of the most comprehensive responsibility who has the conscience for the overall development of man" (BGE 60). For Nietzsche the greatness of the ideal philosopher lies in his "range and multiplicity, in his wholeness in manifoldness"; such a philosopher "would determine rank in accordance with how much and how many things one could bear and take upon himself, how far one could extend his responsibility" (BGE 212). The idea of holding a multiplicity of aspects within a single coherent manifold is, we have seen, the core to Nietzsche's account of what it is to be a person, a sovereign individual capable of exercising free will and responsibility. The idea that Nietzsche's ideal philosopher exhibits both free will and responsibility gets its clearest formulation at the end of section 10 of the third essay of the *Genealogy*;

Is there already enough pride, daring, bravery, self-assuredness in existence today, enough will of spirit, will to responsibility, *freedom of the will* so that henceforth on earth "the philosopher" is truly—possible?

[13] Sue James suggested this nice illustration of the point.

7. NIETZSCHE'S REVISIONARY NATURALISTIC METAPHYSICS OF THE SOUL

That Nietzsche is open to notions of agency free will and agency responsibility that genuinely promote active engagement is further suggested by the analogy with his attitude to the related notion of soul. Generally he is extremely disparaging of the notion of soul because of its Christian baggage. In his notebooks he tells us

> To indulge the fable of "unity", "soul", "person", this we have forbidden: with such hypotheses one only covers up the problem. (KSA 11:577, June–July 1885)

Forbidding the use of these notions fits in easily enough with his general rejection of them. But in the same place he immediately goes on to say, "with such hypotheses one only covers up the problem".

But why does he talk here of covering up a problem? What is the problem if all such things are simply irredeemable fictions? The problem being covered up is that unity, soul, personhood are not pre-given existences but rare achievements to be gained by hard effort. The fable is in the notion that these have already been achieved, that we as mere humans already have unity, are persons.

In BGE 12 Nietzsche hints that his more perceptive reader will see beyond the clumsy naturalists simple rejection of the soul with its associated notion of superlative metaphysical free will and realize that an imminent notion of soul is now available:

> Between ourselves, it is not at all necessary to get rid of "the soul" at the same time, and thus to renounce one of the most ancient and venerable hypotheses—as happens frequently to clumsy naturalists who can hardly touch on "the soul" without immediately losing it. But the way is now open for new versions of the soul-hypothesis; and such conceptions as "mortal souls" and "soul as subjective multiplicity" and "soul as social structure of the drives and affects" are henceforth to have citizen rights in science.

Now one might be tempted to simply treat this suggestion of a new notion of soul, and the characterization of the sovereign individual as a being with agency free will, as isolated incidents in Nietzsche's wider corpus. But against this it is first worth noting that there are other places where Nietzsche uses free will in just this positive way. In *The Gay Science* he writes,

> Once a human being reaches the fundamental conviction that he must be commanded, he becomes a "believer". Conversely one could conceive of such a pleasure and power of self-determination, such a freedom of the will that the spirit would take leave of all faith and every wish for certainty, being practiced in maintaining himself on insubstantial ropes and possibilities and dancing ever near abysses. Such a spirit would be the free spirit par excellence. (GS 348)

Indeed the idea of freedom through self-determination is central to his characterization of his greatest hero Goethe,

> What he wanted was totality . . . he disciplined himself to wholeness, he created himself . . . . Goethe conceived of a person, a strong, highly educated, accomplished in all corporeal matters, self-controlled, self respecting, who can dare to allow himself the whole range and richness of naturalness, who is strong enough for this freedom. . . . Such a spirit who has become free stands with a happy and trusting fatalism in the midst of the universe. (TI 49—my translation)

This notion of freedom is echoed in some Nachlass passages, for instance,

> The freest act is that in which our own most strongest most finely practiced nature springs forth, and in such a way that at the same time our intellect shows its directing hand. (KSA 10:258, Spring–Summer 1883)

Nietzsche, then, should not be seen simply as one who rejects received metaphysical notions of free will, autonomy, agency, personhood, and soul, but as one who replaces them with immanent naturalist accounts. Moreover these accounts serve distinctly normative ends; they aim to replace a passive stance and engender a genuinely active creative engagement with the world. Nietzsche aims to change his preferred readers from being mere conduit points of a vast array of conflicting inherited drives into genuinely unified beings. As Nietzsche's Zarathustra puts it,

> And when my eyes flee from the present to the past, it always discovers the same thing: fragments and limbs and dreadful chances—but no men! . . . I walk among men as among fragments of the future: of the future which I scan. And it is my art and aim, to compose into one and bring together what is fragment and riddle and dreadful chance. (Z II 21)[14]

8. NIETZSCHE AND THE MODERN TRADITION ON THE SELF

This interpretation allows us to see a Nietzsche who has come to grips with a central problem of modern philosophy in a way that many of his predecessors, contemporaries, and even successors, have failed to do. If we take part of the central trajectory of modern philosophy to be the move from a religious to a secular worldview, we (should) see that giving up the metaphysics of God and soul raises a crucial problem about exactly what we are. The modern tradition

[14] As Chris Sykes has pointed out to me, the theme that in modern times, or more specifically in modern Germany, as Nietzsche emphasizes in the second of the *Untimely Meditations*, there are no men, but only fragments of men is clearly a deliberate echo of the same charge that occurs in Hölderlin's *Hyperion*:

> I can think of no people more at odds with themselves than the Germans. You see artisans, but no men, thinkers, but no men . . . is this not like a battlefield on which hacked-off hands and arms and every other limb lie scattered about. (Hölderlin 1952: 420—my translation)

offers a number of answers; we are in essence reasoners (Descartes, Kant)[15]; we are bundles of sensations (Hume). None of these answers are particularly satisfactory. Nietzsche offers an interesting and rather original alternative. He claims that in a sense we do not exist.[16] This is not a version of that kind of academic, philosophic, scepticism that brings philosophy into deserved disrepute. The existence of human bodies, like the existence of the so-called external world is not something Nietzsche would ever dream of really denying. What Nietzsche questions is whether there are genuine selves inhabiting these bodies. In place of empiricist or rationalist accounts of the self, Nietzsche offers, what might be called, a naturalist-aestheticist account: To have a genuine self is to have an enduring coordinated hierarchy of drives. Most humans fail to have such a hierarchy; hence they are not sovereign individuals. Rather they are a jumble of drives with no coherent order. Hence they are not genuine individuals or, we might say, selves. As Nietzsche says with particular reference to modern scholars

> Whatever still remains in him of a "person" strikes him as accidental, often arbitrary, still more disturbing; to such an extent he has become a passageway and reflection of strange forms and events even to himself. (BGE 207)

Nietzsche's various attacks on the Kantian notions of autonomy and free will have multiple objectives. The negative objectives are to show that the notion of a will that transcends the causal order is intellectually unacceptable, and to attack the notion of deserts—objectives hardly unique to Nietzsche. The positive, and more profound and original, objective is to offer his readers the challenging notion that agency free will, genuine autonomy, and hence existence as an individual and self, is possible for some. This challenge should awaken his readers to the profoundly disturbing possibility that they themselves are not yet persons.

9. THE NORMATIVE DIMENSION TO NIETZSCHE'S ACCOUNT OF AGENCY FREE WILL

Now one might object that the concept of agency free will that we have ascribed to Nietzsche lacks a normative dimension that is intrinsic to any notion of free will. Acts involving deserts free will are typically seen as those which should be punished/rewarded. But acts involving agency free will, especially for those

[15] Of course, Descartes and Kant were theists. Nevertheless, the intellectual trajectory of their work, whether intended or not, was inevitably towards a secular worldview.

[16] The Hume of the Appendix to his *Treatise* concedes that the account of the self given in the *Treatise* rendered the self non-existent (Cf. Hume 1978: 635–6). However this for Hume was an unintended and unacceptable consequence of his philosophy. For Nietzsche our general failure to achieve genuine selfhood, and the possibility, for at least a select few, of overcoming that failure, is the very point of much of his philosophy.

who deny deserts free will, might be seen as carrying no normative dimension. The worry here is that the terminology of agency free will amounts to a clever re-description that fails to capture the normative dimension that is essential to any free will debate. But this merely fails to identify the locus of the normativity relevant to agency free will. Deserts free will says that one must have acted with free will if one's action is to merit punishment or reward, so the normative dimension flows from, among other things, the fact that the act involved deserts free will. In the case of agency free will the normativitity is in the actual imperative to exercise agency free will, to be an agent rather than a mere cog in the causal network. On could of course endorse the possibility of agency free will and be indifferent as to whether anyone ever achieved agency free will. But in fact those, like Kant and Nietzsche, who endorse agency free will typically see it as something valuable. Indeed, if Nietzsche's account is correct, in the first person case it is hard to see how one can be indifferent to whether one has free will since that is tantamount to be indifferent as to whether one really exists. That Nietzsche wishes to promote the development of genuine individuals, that, like Kant, Nietzsche sets autonomy as a goal, clearly indicates that he endorses agency free will as an aspiration.[17]

10. A CONCLUDING NOTE ABOUT THE PHENOMENOLOGY OF FREE WILL

The account of free will offered above on Nietzsche's behalf may be correctly termed a third person, as opposed to a first person, type account. It is not a first person account in two related senses. First, it does not seek to explain the phenomenology of free will, what if feels like when one feels one is freely choosing to do A rather than B. Second, it does not place the subjective I at the core of the account. Recall, according to the account given above an act is free when it is an expression of the character from which it originated. Now character was basically

[17] This is not to say that Nietzsche thinks that only through conscious acceptance of the goal of achieving a coherent self can we realize this goal, and that he intends his works to influence us, or at least his better readers, towards such a conscious acceptance of this goal. Alternatively, one might argue that Nietzsche intends his work to somehow affect us at a level below full conscious apprehension, so that our drives are aroused more through emotional responses to his text than through conscious cognitive apprehension of propositional contents. It is worth noting, for instance, that Nietzsche says of the three essays of *On the Genealogy of Morals*

> Regarding expression, intention, and the art of surprise, the three inquiries, which constitute this *Genealogy,* are perhaps uncannier than anything else written so far. [EH, *Genealogy of Morals a Polemic*]

In the very first section of the preface of *On the Genealogy of Morals* Nietzsche paints a picture of a dumbstruck audience who ask "utterly surprised and disconcerted, 'What was that that we have just experienced?' " This, arguably, adds some weight to the idea that Nietzsche is aiming at a kind of therapeutic effect through influencing us at some level below consciousness.

glossed in terms of a coherent, ordered, hierarchy of drives. This account, unlike most traditional ones, does not place a conscious I, located separately from the drives, at the centre of the account.

Regarding the phenomenology of free will, I take Nietzsche to be largely dismissive of its import. That on various occasions one feels one is freely choosing A over B is for Nietzsche not in itself indicative of any genuine freedom. Rather he treats such feelings as something to be explained away. He often glosses it in terms of one drive defeating another which is then followed by a conscious identification with the winning drive (see D109, BGE 19).[18] More generally Nietzsche is on the whole fairly dismissive of the import of consciousness:

> Leibniz's incomparable insight that has been vindicated not only against Descartes but against everybody who had philosophized before him—that consciousness is merely an accident of experience and not its necessary and essential attribute. (GS 305)[19]

Regarding the notion of an I separate from the drives, this for Nietzsche would be an empty mere geometric point. For Nietzsche there is no such I.[20] Consciousness, on this view, is to be regarded as the by-product of various drives, much as Freud saw the ego as a derivative of the original unitary id, and that consciousness may make the faulty assumption that it in itself constitutes a genuine unity, a self-standing I separate from the drives. But any such assumption is false.[21]

REFERENCES

WRITINGS BY NIETZSCHE

The Antichrist, in *Twilight of the Idols and The Antichrist*, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1968).

Beyond Good and Evil, trans. Walter Kaufmannn (New York: Vintage, 1966).

Ecce Homo in *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmannn and R.J. Hollingdale (New York: Vintage, 1969).

[18] Here I am essentially in agreement with Leiter (this volume).

[19] Indeed, Nietzsche often claims that conscious willing is largely epiphenomenal, and sometimes seriously flirts with the conclusion that all conscious phenomena are totally epiphenomenal:

> Finally, why could a "purpose" not be an *epiphenomenon* (*Begleiterscheinung*) in the series of changes that brings about the purposive action—a pale image sketched in consciousness that serves to orient us concerning events, even as a symptom of events, *not* as their cause? . . . Are not all phenomena of consciousness merely terminal phenomena, final links in a chain? (KSA 12:247)

[20] For more on Nietzsche's denial of the I see Sebastian Gardner's paper in this volume, especially his first two footnotes which give detailed listings of passages, mainly from Nietzsche's notebooks, which elaborates this theme. Gardner is not at all sympathetic to Nietzsche's denial of the I.

[21] Thanks are due to Chris Janaway, Sebastian Gardner, Chris Sykes and Gudrun von Tevenar for comments on early drafts of this paper, and to participants of the Gemes/Leiter Nietzsche Seminar held at the University of London in the autumn term of 2005–6, and to participants in the April 7–9, 2006 conference on Nietzsche on Self, Agency, and Autonomy, for helpful discussions on later versions of this paper.

The Gay Science, trans. Walter Kaufmann (New York: Vintage, 1974).
On the Genealogy of Morals, in *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage, 1969).
Human, All Too Human, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1986).
Sämtliche Werke: Kritische Studienausgabe in 15 Einzelbänden, ed. G. Colli and M. Montinari (Berlin: de Gruyter, 1988).
Thus Spoke Zarathustra, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1975).
Twilight of the Idols, trans. R. J. Hollingdale, in *Twilight of the Idols and The Antichrist*, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1968).
Untimely Meditations, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1986).

OTHER WORKS CITED

Gemes, Ken (2001). 'Post-Modernism's Use and Abuse of Nietzsche'. *Philosophy and Phenomenological Research*, 62, 337–360.
____(2006). 'We Remain of Necessity Strangers to Ourselves: The key message of Nietzsche's *Genealogy*', in Christa Acampora (ed.), *Nietzsche's Genealogy of Morals: Critical Essays* (Lanham, MD: Rowman & Littlefield).
Hölderlin, Friedrich (1952). *Dichtungen und Briefe* (München: Winkler-Verlag).
Hume, David (1978). *A Treatise on Human Nature*, edited, with an analytical index by L. A. Selby-Bigge (Oxford: Clarendon Press).
Leiter, Brian (2003) *Nietzsche on Morality* (London: Routledge).
Watson, Gary (1996). 'Two Faces of Responsibility', *Philosophical Topics*, 24: 227–48.

This page intentionally left blank

3

Autonomy, Affect, and the Self in Nietzsche's Project of Genealogy

Christopher Janaway

I

Nietzsche is well known for stating that there is 'only a perspectival "knowing"'. What has been less remarked is the extent to which he thereby stands in radical opposition to a common philosophical position concerning the relationship between knowledge and the affects. On this issue we may let Schopenhauer speak for the philosophical mainstream of many centuries:

> In order to see that a purely objective, and therefore correct, apprehension of things is possible only when we consider them without any personal participation in them, and thus under the complete silence of the will, let us picture to ourselves how much every affect [*Affekt*] or passion [*Leidenschaft*] obscures and falsifies knowledge, in fact how every inclination or aversion [*Neigung oder Abneigung*] twists, colours, and distorts not merely the judgement, but even the original perception of things.[1]

Yet this is exactly what Nietzsche rejects in *On the Genealogy of Morality* III 12, whose chief conclusions are delivered in the following statement:

> There is *only* a perspectival seeing, *only* a perspectival 'knowing'; and *the more* affects we allow to speak about a matter, *the more* eyes, different eyes, we know how to bring to bear on one and the same matter, that much more complete will our 'concept' of this matter, our 'objectivity' be.

A passage in the same section glosses Nietzsche's new understanding of 'objectivity' as:

> the capacity to have one's pro and contra *in one's power*, and to shift them in and out: so that one knows how to make precisely the *difference* in perspectives and affective interpretations useful for knowledge.

[1] Schopenhauer (1969: ii. 373), translation slightly modified from Payne's to bring out the similarity to Nietzsche's vocabulary.

What is it that makes knowing necessarily perspectival? What is it about the condition of the knower that corresponds to the in-built constraints of specificity of place and function that make all seeing perspectival? Nietzsche actually says rather little about this in our passage. He mentions the 'active and interpretive forces', which presumably pertain not just to the 'eye' but to that which it is a metaphor for, i.e. the knowing human mind. Nietzsche perhaps trades on the post-Kantian platitude that knowledge of a world of objects has as a condition that the human mind actively contribute organizing structures to any data it receives, so that non-selective passive reception can be no model of knowledge at all. However, Nietzsche's 'active and interpretive forces' are not very much like Kant's synthesis of intuitions under concepts of the understanding and in some respects more akin to Schopenhauer's notion that the intellect's operations are in the service of the will. The heavy emphasis in our passage falls, not on purely cognitive Kant-style features of the mind, but on *affects*. It is the affects—the very mental states that for the philosophical orthodoxy 'twist, colour, and distort' judgement and perception—that Nietzsche portrays as enabling and expanding knowledge. These points are hammered home in the last half page of section 12: 'To eliminate the will altogether, to disconnect the affects' would be to disable knowledge; 'To have one's pro and contra in one's power' is to make one's knowledge more 'objective'; the plurality of affects, the greatest possible difference in affective interpretations, is 'useful' for knowledge and makes it more 'complete'.

What is an affect? At times Nietzsche talks simply of 'inclinations and aversions', 'pro and contra', or 'for and against'. It seems that all affects are at bottom inclinations or aversions of some kind. But their range is extensive. In the *Genealogy* and *Beyond Good and Evil* alone he explicitly uses the term for the following: anger, fear, love, hatred, hope, envy, revenge, lust, jealousy, irascibility, exuberance, calmness, self-satisfaction, self-humiliation, self-crucifixion, power-lust, greed, suspicion, malice, cruelty, contempt, despair, triumph, feeling of looking down on, feeling of a superior glance towards others, desire to justify oneself in the eyes of others, demand for respect, feelings of laziness, feeling of a command, and brooding over bad deeds.[2] Affects are, at the very least, ways in which we *feel*. Many specific instances are what we would call emotions, some are perhaps moods, while affects of 'commanding' or of 'looking down on' someone are not obviously describable as either moods or emotions. The class of affects is likely to include further felt states such as an instinctual like or dislike for something, a sense of unease, a faint thrill at a certain thought, and so on until we reach states for which we have no terminology:

> words really exist only for *superlative* degrees of these [inner] processes and drives. [. . .] Anger, hatred, love, pity, desire, knowledge, joy, pain—all are names for *extreme* states:

[2] See BGE 19, 23, 187, 192, 260; GM I 10, 13; GM II 11; GM III 15, 20.

the milder, middle degrees, not to speak of the lower degrees which are continually in play, elude us, and yet it is they which weave the web of our character and destiny. (D 115)

Some affects are beneath accurate apprehension by ourselves, and some are unconscious.[3] But all seem to be *feelings* of one sort or another. And if we respect the fact that Nietzsche gives such prominence to affects in his discussion of perspectival knowing, we shall have to surmise that for him the in-built constraint upon knowledge that makes it 'only perspectival' lies in the knowing subject's affective nature.

What, then, can Nietzsche mean when he talks of the affects enabling and enhancing knowledge? I suggest that we begin not with a generalization about all knowing, but with a specific example extremely close to home: Nietzsche's own interpretive task in the *Genealogy*, his search for knowledge about the various phenomena of morality. I have argued elsewhere[4] that arousing the affects is central to Nietzsche's aims as a writer, and that it deserves to be so because of the prominence of affects in his explanation of the genesis of our moral attitudes. For example, in order to further our understanding of the slave-revolt in morality Nietzsche encourages us to recognize our own ambivalent inclinations and aversions—our mixed feelings for or against compassion, aggression, humility, prowess, equality, nobility—and to reconstruct the history of attitudes to 'good' and 'evil' in imaginative engagement with the feelings of both the oppressed slave and the self-defining master. When he invites us to identify with the fears and resentment of a defenceless sheep and with the fond disdain of a natural predator, he provides a way of gaining insight into the origins of the modern notion of responsibility for action. Again, in convincing us of cruelty's role in the genesis of guilt and punishment, he seduces us into acknowledging, beneath our more obvious feelings of anger and disgust, a streak of joyfulness in seeing and making suffer. All these seem good examples of 'making the difference in perspectives and affective interpretations useful for knowledge'—knowledge of the nature of our own moral attitudes. An even more local example occurs in the first half of GM III 12 itself, which invites us to feel the philosophical products of asceticism as a 'triumph', commenting that we should 'not be ungrateful toward such resolute reversals of [. . .] familiar perspectives and valuations'—in other words, to see things through the eyes and feelings of the ascetic is beneficial to our inquiry because it enables us to have a more complete insight into the ideal that, according to Nietzsche, dominates Christian and post-Christian culture.

A further small but significant point about Nietzsche's vocabulary might assist in reorienting our expectations about the kind of knowledge or interpretation Nietzsche has most immediately in mind. He talks of completing our 'objectivity'

[3] See GS 354, and Richardson (1996: 34–8).

[4] See Janaway (2003), Janaway (2006a). Janaway (2007*b*) addresses more fully many of the issues that arise in the present paper—see especially chs. 7 and 12.

and our 'concept' of something—but of what kind of thing? Some accounts of perspectivism foreground the issue of 'knowing an object from a particular perspective',[5] which may carry the suggestion that Nietzsche's main concern is the traditional epistemological problem of knowledge of individual bodies in the external world. But it is unnecessary to read Nietzsche as referring to an *object* in this specific manner.[6] His word here is *eine Sache*, which may mean simply a *matter* or *topic* of discussion. Back in the *Genealogy*'s Second Treatise (section 4) he investigated a particular *düstere Sache*, a gloomy matter, namely 'the consciousness of guilt, the entire "bad conscience" ' which turned out to be an immensely complex set of psychological and cultural phenomena. Another such *düstere Sache* for him is 'reason, seriousness, mastery over the affects, this entire gloomy matter called reflection',[7] and there are similar uses of the term in the *Genealogy* to denote a topic of discussion, an area of human life under investigation.

So Nietzsche's practice in the *Genealogy* suggests the belief that our feeling shocked, embarrassed, disgusted, or attracted by some phenomenon *tells us something about* that phenomenon—that is, that feelings themselves have cognitive potency. Such a construal of perspectival 'knowing' has the advantage of fitting the manner of writing about morality that Nietzsche uses throughout the *Genealogy*. And the otherwise rather obscure notion of 'shifting one's pro and contra in and out' seems an appropriate description of the activity Nietzsche expects of his own more alert and sensitive readers.

So we are faced with two strikingly controversial general claims: (1) That it is impossible for there to be any knowing that is free of all affects, and (2) That multiplying different affects always improves knowing. The second claim could stand even if it were to turn out that some forms of knowledge, such as scientific investigation, for instance, have to be construed as purged of all affects. For Nietzsche can hold that knowing something *only* scientifically gives us a poorer understanding of it than knowing it through a variety of psychological, imaginative, rhetorical means—affect-arousing means—in addition to those of science. His own shift from philological *Wissenschaft* to the manner of writing displayed in the *Genealogy* bears out the intelligibility of such a claim for Nietzsche. It is beyond question that Nietzsche regards the *Genealogy* as providing greater knowledge about morality than any combination of the traditional *Wissenschaften* could have attained unaided.

[5] Leiter (1994: 345).

[6] Cox (1999: 139–68), argues for the stronger claim that Nietzsche's perspectivism cannot be construed as assuming the everyday conception of objecthood, but rather undermines that notion.

[7] GM II 3. Other instances in the *Genealogy*: punishment is the *Sache*, or topic of discussion, in GM II 13; the hatred of rule and hierarchy that dominates science is a *schlechte Sache*, a bad thing (GM II 12); human history itself would have been a *dumme Sache*, a stupid thing, without the spirit that the powerless introduced into it (GM I 7).

Nietzsche's other general claim—that there is no form of knowing that is not affective—looks harder to sustain. We are prompted to seek some theoretical ground for Nietzsche's refusal in principle to allow affect-free knowledge. At this point it will help to consider who or what is the knower, the subject of knowledge, for Nietzsche. Faced with the questions Who knows? Who thinks? Who interprets? Nietzsche's official position is that there is no such subject as ordinarily conceived. He repeatedly urges that we should be suspicious towards the concept of a subject or 'I'. The I is 'just an assumption or opinion, to put it mildly', it has 'become a fairy tale, a fiction, a play on words', and enjoys 'a merely apparent existence'. Instead we are to think of 'soul as subject-multiplicity', and view the self as a plurality of subpersonal elements in competitive interaction with one another, elements that are will-like in character (' "under-wills" or under-souls').[8]

Nietzsche commonly calls such elements 'drives'. In the case of a philosopher, for example, *who he is* is equivalent to 'what order of rank the innermost drives of his nature' stand in, and thinking itself is 'only a relation between these drives' (BGE 6, 36). But it is evident that drives are closely related to affects, for he also says that the social construction that is the self is built out of 'drives and affects' (BGE 12), and talks elsewhere of 'our drives and their for and against'.[9] We may wonder whether drives and affects are even properly distinguishable kinds.[10] It would be foolhardy to expect consistent terminological rigour here, but unless the recurrent expression 'drives and affects' is to be taken as merely pleonastic, we might hypothesize that a drive is a relatively stable tendency to active behaviour of some kind, while an affect, put very roughly, is what it feels like when a drive is active inside oneself. Affects, as we have seen, are glossed as inclinations and aversions or fors and againsts. An affect would then be a positive or negative feeling that occurs in response to the success or failure of a particular drive in its striving, or in response to the confluence of the activities of more than one drive within oneself. In *Human, All Too Human* Nietzsche already thought that 'a drive to something or away from something divorced from a feeling one is desiring the beneficial or avoiding the harmful, a drive without some kind of knowing evaluation of the worth of its objective, does not exist in man'.[11] And in a late notebook entry he asks 'Can we assume a *striving for power* without a sensation of pleasure and unpleasure, i.e. without a feeling of the increase and diminution of power?'[12]

[8] See BGE 12, 16, 17, 19, 34, 54; TI 'Reason in Philosophy' 5; 'The Four Great Errors', 3.

[9] WLN 139 (note 7 [60] (1886–7), previously pub. as WP 481).

[10] As suggested by Cox (1999: 126–7): 'The disposition that composes [subjects] is itself made up of microdispositions—what Nietzsche variously calls "drives" (*Triebe*), "desires" (*Begierden*), "instincts" (*Instinkte*), "powers" (*Mächte*), "forces" (*Kräfte*), "impulses" (*Reize, Impulse*), "passions" (*Leidenschaften*), "feelings" (*Gefühlen*), "affects" (*Affekte*), pathos (*Pathos*), and so on.'

[11] HAH I 32. See also Richardson (1996: 37).

[12] WLN 248 (n. 14 [82] (1888), previously pub. as WP 689).

Given the claim that this plurality of will-like striving components held in tension is what I am, it makes sense prima facie to hold that if I am to *know* anything, it must be through some activity of the drives that compose me and the feelings essentially involved in their activity. There is, officially, nothing else available to do the knowing. Thus in further notebook entries Nietzsche writes: 'It is our needs *which interpret the world*: our drives and their for and against. Every drive is a kind of lust for domination, each has its perspective'; and '*Who interprets?*—Our affects'.[13] Out of the inner tension and competition among striving and feeling dispositions, constantly interpreting and seeking to become master over whatever they encounter, come thinking, meaning, and valuing. It would seem that, given Nietzsche's view of the self, knowing would likewise have to be some aggregate of the multiple affective interpretations that occur within the one human being.

II

Construed in this manner, Nietzsche's conception of perspectival knowledge still faces some hard questions. (1) What makes an aggregate of the activities of someone's drives a case of *knowing*, as opposed to a case of something else? Can the activity of the multiple drives and affects give rise to interpretations that are capable of meeting even the minimum requirement for knowledge, that of being true? (2) If each of the subpersonal elements of the self (whatever they are called) wants mastery, strives, interprets, and grasps something of reality, is not each of them just a miniature subject? Far from abandoning the subject-conception inherited from the philosophical tradition, have we not simply denied it application at the level of the individual human being and smuggled it back into our description of the subpersonal community of drives? (3) How in Nietzsche's conception can one be said 'to have one's pro and contra *in one's power*, and to shift them in out'? How can we 'make [. . .] the difference in perspectives and affective interpretations useful for knowledge'? How can we 'bring to bear' our many affects on one and the same matter, and 'allow [them] to speak' about it? What, in short, must the self be, in order to perform such operations? And finally, (4) can Nietzsche, with his apparently eliminativist conception of the self, explain how the philosopher's conception of the self-conscious subject or 'I', erroneous or not, could even have come about?

Here I shall address only the latter two issues. Our passage from GM III 12 suggests that for 'knowing' to occur, a mere multiplicity of affect-interpretations will indeed not suffice. There must also be those operations upon affects, or upon interpretations, which Nietzsche calls 'having them in one's power', 'shifting

[13] WLN 139, 96 (n. 7 [60] (1886–7), and [2] 190 (1885–6), previously pub. as WP 481, 254).

them in and out', and so on. This then prompts our question (3): what must the self be for these operations to occur? Perhaps some kind of controlling subject? In similar passages in *Beyond Good and Evil* and *Human, All Too Human* Nietzsche speaks blatantly of free agency with respect to one's affects:

> To freely have or not have your affects, your pros and cons, to condescend to them for a few hours; to *seat* yourself on them as you would on a horse or often as you would on an ass:—since you need to know how to use their stupidity as well as you know how to use their fire.[14]

> You shall become master over yourself [. . .] You shall get control over your For and Against and learn how to display first one and then the other in accordance with your higher goal.[15]

It seems clear that there must be some kind of self-conscious unity for the notion of controlling affects and bringing them to bear on one and the same subject-matter to make sense. I have to be aware that affects *A* and *B*, each of which may 'speak' interchangeably about the same subject-matter, are both *mine*. I have to be that which feels both affects, regards itself as feeling them both, and takes some attitude towards its subject-matter in the light now of this affect, and now of that. It seems right to view Nietzsche's suggestion of 'having one's affects within one's control' and manipulating them 'in accordance with a higher goal' as part of his overall revaluative project. In using the fullness of our affective responsiveness to the world, we come to occupy ourselves, as it were, in a more complete and healthy way, to fulfil our potential as cognizers. But if the way in which we are to reach this healthier cognitive state is by rethinking what we are and by conscious identification with as many of our affects as possible, we must arguably be unified self-conscious subjects, subjects of 'I'-thoughts. So we cannot *simply* be a multiplicity of drives and affects, as Nietzsche's official position proclaims.

An alternative, Kantian position would allow for self-conscious unity within experience without reawakening any of the metaphysical notions that Nietzsche repudiates—the self or 'I' as non-empirical, unchanging essence, as thinking substance, as something ontologically primary and irreducible, as cause of thinking. Kant makes the transcendental claim that it is a condition of there being experience at all that it is that of a self-conscious 'I'. But, for Kant, no metaphysical conclusions can be drawn from the holding of this condition, hence his criticisms of rational psychology in the 'Paralogisms of Pure Reason': how this 'I' is realized ontologically is left entirely open by the transcendental claim.[16]

[14] BGE 284, translation modified from Norman's. Norman has 'your stupidity' and 'your fire' for 'ihre Dummheit' and 'ihr Feuer', but the correct translation of 'ihre' and 'ihr' must be 'their'.

[15] HAH I Preface 6 (published 1886). 'For and Against' translates 'Für und Wider', rendered elsewhere as 'pro and con'.

[16] See Kant (1998), 'Transcendental Deduction of the Categories' and 'Paralogisms of Pure Reason'.

If he espoused this Kantian position, Nietzsche could hold that at the level of physiological description there are drives and affects in various sorts of relations, and, compatibly with this, that for there to be the operations of 'controlling' and 'shifting in and out' in regard to the affects, there must be a subject that conceives itself as a single self-conscious 'I'.[17] His eliminativist treatment of the self could then be seen as directed primarily against traditional metaphysical conceptions, and his denial of *any* subject or 'I' regarded as a polemical over-statement.[18]

Given that he eschews this Kantian route, however, Nietzsche's alternative approach towards explaining the (apparent) unity of self-consciousness is the solely naturalistic one of positing integration of the multiple drives under one dominant drive. An example of this can be found in *Daybreak* 109, where Nietzsche gives a penetrating account of ways in which one might 'combat the vehemence of a drive'—Nietzsche gives no specific illustration, but we might imagine someone struggling with overpowering sexual urges or addictive cravings. The discussion concludes thus:

> *that* one *desires* to combat the vehemence of a drive at all, however, does not stand within our own power [. . .] in this entire procedure our intellect is only the blind instrument of *another drive* which is a *rival* of the desire whose vehemence is tormenting us [. . .] while 'we' believe we are complaining about the vehemence of a drive, at bottom it is one drive *which is complaining about another*; that is to say: for us to become aware that we are suffering from the *vehemence* of a drive presupposes the existence of another equally vehement or even more vehement drive, and that a *struggle* is in prospect in which our intellect is going to have to take sides. (D 109)

A single drive can empower itself by subordinating many other drives to its own activity, and Nietzsche sees organization by a dominant drive as giving unity to one's character and actions. That I will to resist my addictive cravings is not 'up to me', is not the resolve of an 'I' that is external to the complex of drives and affects, but is itself the activity of a strong drive within me. There is no controlling self that determines *ex nihilo* what my ends, purposes, and values are. Fair enough. But I have to be, in my own self-conception, a sufficiently unified self that *I* can 'take sides' between the various drives that (though I did not originally will them) I find within myself. Likewise, it is not just that each of the affects I find within myself has a goal of its own, but rather that *I* have a goal in pursuit of which I can flexibly use the affects I feel. When Nietzsche is thinking of his ideal, creatively evaluating, perspectivally knowing individual, he freely imbues this individual with the status of a unified, self-conscious, autonomous subject, in a way that fails to mesh

[17] Even Schopenhauer's more simplistic position could be of some help to Nietzsche here: that we 'find ourselves as' the (unified) subject from our own point of view, while from an objective, third-person standpoint we are nothing but organic functions expressive of will to life (see Schopenhauer (1969: i. 5 and ii. 272–8)).

[18] See Gardner (2008), in the present volume.

comfortably with his eliminativist description of what the individual amounts to 'in reality'.[19]

If it were just that Nietzsche's ideal, value-creating, perspective-wielding 'new philosopher' presupposes the vocabulary of the unified self which his theoretical descriptions deny, the problem could be finessed by saying that Nietzsche sets us a daunting task: out of the base raw material of warring, organic strivings, which is all humanity has the right to assume itself to be, to create something whole, something with form, goal, and concentrated mastery—a recognizable Nietzschean task. The gap between the given and the goal would accord at least with the enormity and rarity of the achievement Nietzsche invokes. But our question (4) shows that this is not the end of the problem. For how could we come to think of ourselves erroneously in the manner of the philosophical tradition, as simple substances, self-transparent rational thinkers, pure subjects of knowledge, radically free, neutral subjects of choice, and so on? How did we get to regard ourselves as unitary selves at all, erroneously or not? As Sebastian Gardner succinctly puts it,

> How, except in the perspective of an I, of something that takes itself to have unity of the self's sort, can a conception of unity sufficient to account for the fiction of the I be formed? (As it might be put: How can the 'idea' of the I *occur* to a unit of will to power or composite thereof—or to anything *less* than an I?)[20]

In other words, if we were not already unitary, self-conscious selves, how could we have imagined that we were? This raises the prospect that Nietzsche's eliminativist picture of the self may be out of step not only with his project of re-evaluation and his conception of enhanced knowing, but also with his diagnosis of the origins of our metaphysical errors. If only a unified self can make these metaphysical errors, and only a unified self can have the goals and perspectival adaptability that lead to healthier knowing and valuing, then, though we can learn not to think of ourselves as pure metaphysical subjects, Nietzsche's philosophy as a whole demands that we do not regard ourselves *only* as complex hierarchies of drives and affects.

In various ways, then, Nietzsche appears to require a unified self of a kind that his official position would deny. It is hard to find a resolution to this predicament. But some of Nietzsche's remarks—that human beings are a combination of 'creature' and 'creator' in a way that is difficult to comprehend (BGE 225), and that there is no need to follow clumsy naturalists in abandoning the hypothesis of 'the soul' (BGE 12)—show him alive to a central tension in his view of the self, and even keen to cultivate our awareness of that tension. I would argue that his perspectivism is another case in point: in support of

[19] I am indebted to Gardner's presentation of this 'lack of fit' argument.

[20] Gardner (2008: 000), in this volume. Some while ago I expressed the problem similarly: 'Can a collection of sub-personal drives fabricate a unitary self that comes to regard those drives as its own? Or must there be a presupposed unitary self as author of the fiction?' (Janaway (1989: 355)).

the view that our interpretations are saturated and constituted by a plurality of feelings he dissolves the self into a multiplicity of affects and drives. But his aims of improving our capacity for knowing and skilfully using our affects demand more of a self than that: he needs his enquirer to be an active and sufficiently unified self that can represent its subject-matter truly, that rides on top of the inner multiplicity, and that can self-consciously adopt attitudes towards it.

III

The vexed question of free will raises similar issues. Nietzsche talks of free will towards the end of the First Treatise of the *Genealogy* and towards the beginning of the Second, and may appear to contradict himself within the space of a few pages. He describes as nonsensical and false the belief in a 'neutral "subject" with free choice' standing behind an individual's actions (GM I 13); yet he says there has existed or may exist a type of individual who 'has become free' and is a 'lord of the *free* will' (GM II 2).[21] The text leaves us uncertain as to who this sovereign individual is, was, or might be. He or she is described as an end-product of the conformist 'morality of custom', a mode of evaluation prior to the Christian morality Nietzsche is out to re-evaluate in the *Genealogy*.[22] But are 'sovereign individuals' supposed to have existed after the age of the morality of custom was over or during its later stages? And are they supposed to have existed once and then faded away into history, or are there sovereign individuals around today? Or have they never existed? The tone suggests that Nietzsche may be describing an ideal type, giving us what Aaron Ridley has called 'a sort of foretaste of the (enlightened) conscience of the future'.[23] Many such questions are left open. But whatever else the mysterious 'sovereign individual' is, he is not supposed to be a nonsense or a falsehood. Nietzsche must regard the sovereign individual's achievement of freedom as something other than his becoming a neutral subject with free choice.

It has been claimed that being a sovereign individual is for Nietzsche constitutive of being truly human.[24] But this is difficult to support. For although Nietzsche attributes to the sovereign individual a 'feeling of the completion of man himself' (GM II 2), he emphasizes the distinction and superiority of the sovereign individual over other types of human individuals who lack power, pride, and autonomy.[25] In the sense of free will at issue here not every human

[21] A distinction made by Ken Gemes between deserts free will and agency free will is an appealing way of resolving the tension—see Gemes (2009). Janaway (2006*b*), addresses similar issues, and overlaps to some extent with Parts III and IV of the present paper.

[22] Nietzsche introduces the 'morality of custom' (*Sittlichkeit der Sitte*) in D 9.

[23] Ridley (1998: 18). [24] See Havas (2000: 94–5).

[25] The point is made in reply to Havas by Ridley (2000: 106–7).

being will have free will, or at least not to the same degree.[26] Nietzsche here ignores the global, metaphysical question whether absence of necessity is possible in human agency, and poses a cultural and psychological question about qualities and conditions that mark out certain human beings as peculiarly admirable or valuable. Free will in this sense is a variably realizable condition, not a universal one. It is an achievement, or a blessing, of the few, and can occur only in some cultural circumstances with people of certain character types.

A further pointer in this direction appears right at the end of the Second Treatise, where Nietzsche envisages a different creative kind of spirit, a rare and exceptionally strong 'human of the future', a 'bell-stroke of noon and of the great decision, that makes the will free again, that gives back to the earth its goal and to man his hope; this anti-Christ and anti-nihilist; this conqueror of God and of nothingness' (GM II 24). Such talk of making the will free *again* suggests a fall and redemption pattern: at some time in the past, as a product of the harshly repressive 'morality of custom', there became possible sovereign individuals with the characteristic quality of having a free will. Since that time the post-Christian morality of selflessness has been victorious, positing the desirability of guilt and self-suppression and the conception of the non-self-suppressing individual as blameworthy for not making the supposedly available choice to be harmless. In some future we might cast off this conception of morality, and the will could be free again.

The individual with free will contrasts starkly with the morality of custom (*die Sittlichkeit der Sitte*) because, as Nietzsche provocatively puts it, ' "autonomous" and "moral" are mutually exclusive' (GM II 2). For there to be values at all, Nietzsche suggests, there had to be a long prehistory in which simple conformity to tradition determined what was good, departure from tradition what was bad and fit to be curbed. Civilization begins with the proposition 'any custom is better than no custom', and tradition is a 'higher authority that one obeys [. . .] because it *commands*' (D 16, 9). Yet the end-product or 'fruit' of this whole constraining process is an individual 'resembling only himself', having the capacity to be 'free again from the morality of custom', to have an 'independent [. . .] will' and be 'autonomous'. Nietzsche says much in a short space here, perhaps grasping for a vocabulary that will capture his insights. The sovereign individual's will is 'free', 'his own', 'independent', 'long', and 'unbreakable'; and in virtue of this will the sovereign individual is permitted to promise, has 'mastery over himself', has his own standard of value, is permitted to say 'yes' to himself,[27] and has a consciousness of his 'superiority' and 'completion'. To be *permitted* to make promises, one must not only be minimally capable of promising but have the power to fulfil one's promises and the integrity to promise only what one genuinely has the will to do. This suggests a kind of self-knowledge in which one

[26] This is a central point in Gemes (2009).

[27] GM II 3. All immediately surrounding quotations and paraphrases are from GM II 2.

is properly conscious of what it is that one wills, and confident of the consistency with which one's will is going to maintain itself intact until the moment at which it can be delivered upon. The sovereign individual can count upon himself to act consistently, to be the same in the future when the time comes to produce what he promised in the past. Understanding oneself in this way, one will presumably attain a justified sense of satisfaction in one's power and integrity, and value others, not according to their conformity to some general practice imposed from without, but according to their manifestation of the kind of power and integrity one recognizes in oneself.

This positive conception of free will, then, involves acting fully within one's character, knowing its limits and capabilities and valuing oneself for what one is rather than for one's conformity to an external standard or to what one ought to be. In the later *Twilight of the Idols* Nietzsche eulogizes Goethe as 'a spirit *become free* [*freigewordner*]', who 'dares to allow himself the whole compass and wealth of naturalness, who is strong enough for this freedom' and 'stands in the midst of the universe with a joyful and trusting fatalism'.[28] One becomes free in accepting and affirming oneself as a whole, and rather than seeing the necessity or fatedness of one's character as an inhibition or obstacle to action, one sees it as the condition of and opportunity for true self-expression.

We might be able to conceive of something like the following as an approximation to Nietzsche's sovereign individual: someone who is conscious of the strength and consistency of his or her own character over time; who creatively affirms and embraces him- or herself as valuable, and who values his or her actions because of the degree to which they are in character; who welcomes the limitation and discipline of internal and external nature as the true conditions of action and creation, but whose evaluations arise from a sense of who he or she is, rather than from conformity to some external or generic code of values. This is a glimpse of the sense in which free will might be attained, or regained for Nietzsche.

Free will, thus adumbrated, needs to be kept distinct from the notion of the 'neutral "subject" with free choice' conceived as underlying our actions, and from the notion of the self as *causa sui*, or uncaused cause of itself. In *Beyond Good and Evil* 21, Nietzsche unequivocally states that the notion of being self-caused is a contradiction. However, that passage contains no premiss that resembles 'there can be free will only if there is a *causa sui*'.[29] Here Nietzsche is not even pursuing the question whether there is or is not free will. Rather he is at his usual genealogical business: flushing out an underlying affective state—'*the longing for* "freedom of the will" in the superlative metaphysical sense'—and hypothesizing an explanation for its genesis and persistence. How do we come to have a thirst

[28] TI 'Expeditions of an Untimely Man', 49. My translation of *freigewordner*, in preference to Hollingdale's 'emancipated'.

[29] As against the reading of Leiter (2002: 88–90). I am helped here by Owen and Ridley (2003).

for this extremity of metaphysics, a self that makes itself out of nothing, and why is this extreme view lodged so firmly in the modern consciousness? Nietzsche's answer here is that we cannot stomach any sense that we are not *wholly* in control of ourselves. Nietzsche is not doing metaphysics, rather unearthing the valuations of ourselves that underlie our inclining to a certain metaphysical position.[30]

So far this is, admittedly, compatible with Nietzsche's rejecting free will altogether. But Nietzsche next asks his reader 'to rid his mind of the reversal of this misconceived concept of "free will": I mean the "un-free will" [. . .] The "un-free will" is mythology'. There is reason to believe, Nietzsche now argues, that in nature itself there is no 'causal association', 'necessity', or 'psychological un-freedom'; we merely project such notions on to reality. When he asks why we make this projection, he is again seeking psychological explanations for some feelings that lie beneath our thoughts. But he is equally suspicious of the antithesis of this extreme metaphysical notion:

> It is almost always a symptom of what is lacking in a thinker when he senses some compulsion, need, having-to-follow, pressure, un-freedom in every 'causal connection' and 'psychological necessity'. It is very telling to feel this way—the person tells on himself. And in general, if I have observed correctly, 'un-freedom of the will' is regarded as a problem by two completely opposed parties, but always in a profoundly *personal* manner. The one party would never dream of relinquishing their 'responsibility', a belief in *themselves*, a personal right to *their own* merit (the vain races belong to this group—). Those in the other party, on the contrary, do not want to be responsible for anything or to be guilty of anything; driven by an inner self-contempt, they long to be able to *shift the blame* for themselves to something else. (BGE 21)

In his earlier *Human, All Too Human*, while still influenced by Paul Rée's straightforward determinism (and to come extent by Schopenhauer's similar position),[31] Nietzsche had consistently described the belief in free will as an error and referred to the *unfreedom* of the will as 'total' and 'unconditional'.[32] Human beings, he claimed, are no more free than animals (HAH I 102), or indeed than a waterfall in which we may 'think we see [. . .] capriciousness and freedom of will' (HAH I 106). In contrast, when he approaches the neutral 'subject' and *causa sui* in the *Genealogy* and *Beyond Good and Evil*, Nietzsche's interest is genealogical: he is concerned less with resolving the traditional metaphysical debate over the question whether one's actions could have turned out otherwise under identical causal conditions, and more with diagnosing the affective origins of metaphysical beliefs at both extremes of the debate. Nietzsche the genealogist asks what affective psychological states explain the origination of these extreme

[30] Other metaphysical beliefs receive analogous treatment by Nietzsche. For instance, we do not find him arguing against the existence of God; instead he asks what affective states explain our attachment to belief in God.

[31] For discussion of these influences, see Janaway (2006*b*) or Janaway (2007*b*), ch. 7.

[32] See HAH I 18, 39, 99, 102, 106; HAH II/1, 50, 33; HAH II/2, 12.

metaphysical pictures of ourselves. Similarly in *Genealogy* I 13, he asks for the psychological origins of belief in the indifferent subject unconstrained by nature and circumstances, and finds the answer in *ressentiment* and its outgrowth, the felt need for a target for blame.

To gain a glimpse of the 'free will' that may pertain to the sovereign individual, then, we must step outside our learned moralistic preoccupation with blame and with the neutralizing of character differences in explaining action, and look beyond the dichotomy between the notion of *causa sui* or radical independence from nature and the 'total unfreedom' of Nietzsche's earlier position: a dichotomy between two myths, as Nietzsche has warned us, myths that prevail because they are driven by differing affective impulses within us.

IV

Nietzsche is hoping for a revaluation of values: he wants some of us at least to change our allegiance away from the values of selflessness as he has diagnosed them, and to regard them as symptoms of sickness and decline, which we will do our best to distance ourselves from in future. But what is this change in allegiance, and how might it occur? An important aspect to revaluation is the claiming of values as one's own. Rather than adhering to values which are received, traditional, generic, universal, one is to discover one's own personal values. For example, Nietzsche says:

> Let us [. . .] *limit* ourselves to the purification of our opinions and value judgements and to the *creation of tables of what is good that are new and all our own*: let us stop brooding over the 'moral value of our actions'! Yes, my friends, it is time to feel nauseous about some people's moral chatter about others. Sitting in moral judgement should offend our taste. (GS 335)

The strictly genealogical project of discovering truths about the psychological origins of our evaluations does not as such revalue them, but, as Nietzsche says in a notebook entry,[33] 'for our feelings [. . .] [it] reduces the value of the thing which originated that way, and prepares a critical mood and attitude towards it'. He also puts this by saying 'Your insight into *how such things as moral judgements could ever have come into existence* would spoil these emotional words [such as "duty" and "conscience"] for you' (GS 335). The effect, I take it, could also be described as a loss of one's more or less automatic emotional alignment with received values, a suspension of the single-dimensional 'pro and contra' inherited from the Christian culture of Nietzsche's most typical readers. This suspension allows a space for a new evaluation and a shift or reversal in values, which Nietzsche often describes in ways which seem to presuppose agency, judgement, and choice.

[33] WLN 95 (n. 2 [189] (1885–6), previously pub. as WP 254).

In *Daybreak* Nietzsche describes the change in valuation he seeks with the phrase 'we have to *learn to think differently*'—i.e. outside the moral evaluative oppositions of good and evil, egoistic and selfless—'in order at last, perhaps very late on, to attain even more: *to feel differently*' (D 103). Also in *Daybreak* Nietzsche adumbrates a kind of liberation of the individual's thinking, not from feelings *per se*, which for him would be impossible and undesirable,[34] but from feelings which are not original or appropriate to the individual. Instead of carrying around 'valuations of things that originate in the passions and loves of former centuries' (GS 57) and giving 'obedience to one's grandfather and grandmother and their grandparents', one is to honour what he strikingly calls 'the gods which are in *us*: our reason and our experience' (D 35). Other passages suggest that a revaluation of values will be an act of placing trust in values that are authentically one's own, an autonomous decision taken in the light of self-understanding.

> It is selfish to consider one's own judgement a universal law, and this selfishness is blind, petty, and simple because it shows that *you haven't yet discovered yourself or created for yourself an ideal of your very own* [my emphasis] [. . .] No one who judges, 'in this case everyone would have to act like this' has yet taken five steps towards self-knowledge. [. . .] We, however, want to *become who we are*—human beings who are new, unique, incomparable, who give themselves laws, who create themselves! To that end we must become the best students and discoverers of everything lawful and necessary in the world. (GS 335)

Nietzsche here predicates both self-knowledge and autonomy of those who would successfully follow him. Against this one may range many passages in which Nietzsche asserts that self and will are illusions, that there is no internal 'helmsman' controlling one's actions, no unitary subject of thought or action known to oneself by privileged access, only a subterranean multiplicity of competing or hierarchically organized drives, of which one's knowledge will always be incomplete.[35] However, the fact that self-knowledge (likewise selfhood as such, as Gemes shows[36]) is hard, that most have very little of it, and even that no one ever attains it completely—none of this shows that self-knowledge is impossible: only that it is rare among human beings, that it is a task set for a few of us rather than a given, and that its achievement is a matter of degree.

A relevant example comes at the end of the *Genealogy*'s Second Treatise. Here Nietzsche takes himself to have shown that guilt came to be regarded as a good in the Christian worldview because the conception of our natural instinctual selves as an ultimate transgression against God allowed us the most powerful guarantee

[34] As witness GM III 12 discussed above.

[35] See D 109, 116, 119, 129; GS 335, 360; BGE 16, 17, 19, 34, 54; TI 'Reason in Philosophy' 5, 'The four Great Errors' 3.

[36] Gemes (2009).

of being able to vent our in-built drive towards cruelty upon ourselves. Nietzsche evaluates this state of self-torture as 'the most terrible sickness that has thus far raged in man' (GM II 22). But then he offers us the healthy alternative, 'a reverse attempt [. . .] namely to wed to bad conscience the *unnatural* inclinations, all those aspirations to the beyond, to that which is contrary to the senses, contrary to the instincts, contrary to nature' (GM II 24)—though he doubts that any but that most exceptional human being of the future, the redeeming, creative spirit of *great health*, will be able to accomplish this.

How to characterize the change of allegiance in values that Nietzsche here imagines someone undergoing? We start with the observation that people feel guilt and regard life lived with an enduring guilty conscience as having positive value. We offer to explain these phenomena in terms of historical psychological states: in brief, an instinctual drive to inflict cruelty, internalization of the instincts, and rationalization of self-cruelty by the invention of a theistic metaphysics.[37] We judge this psychological complex a sickness, allowing ourselves, as Nietzsche says, to feel horrified and unnerved by the sadness of it (GM II 22). Although we are the inheritors of Christian attitudes of disapproval to what is labelled egoistic, we can take a step back from our accustomed valuations, and then—if strong enough—try to bring ourselves to feel negatively towards ourselves if we experience any continuing temptation to despise our natural instincts and inclinations, or to hope for a higher, otherworldly order of values. So the process of reversal Nietzsche envisages is cognitive at many stages. One *comes to believe* a certain explanation as true, one *judges* a set of psychological states as unhealthy, one *tries* to feel a new set of affects, and *identifies oneself with* specific critical second-order attitudes regarding certain of one's feelings.

Nietzsche imagines that by examining our own deep-seated attitudes of inclination and aversion, by accepting hypotheses about their origin in past psychological configurations such as those of the ancient masters and slaves, by reflecting on which values we feel as most congenial to our characters, we may attach ourselves to a new set of values. The latter step of becoming free from the inherited values of morality requires, I argue, the conception of oneself as deciding, choosing, and trying as a genuine agent. Such genuine agency does not require that one be a neutral subject of free will that has unlimited possibility of action unconstrained by character and the causal order. In that sense there is no free will. But it does require, as Nietzsche says, that we rid ourselves of the other myth, that of the total unfreedom of the will. So it is wrong to think that Nietzsche wishes to exclude creative agency from his picture of humanity, because without it his proposed critique of moral values and his project of learning to think and feel in healthier ways would make little sense.

[37] For a fuller account see Janaway (2007*a*) and Janaway (2007*b*), ch. 8.

REFERENCES

WRITINGS BY NIETZSCHE

Beyond Good and Evil, ed. Rolf-Peter Horstmann and Judith Norman, trans. Judith Norman (Cambridge: Cambridge University Press, 2002).

Daybreak, ed. Maudemarie Clark and Brian Leiter, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1997).

On the Genealogy of Morality, trans. Maudemarie Clark and Alan J. Swensen (Indianapolis: Hackett, 1998).

The Gay Science, ed. Bernard Williams, trans. Josephine Nauckhoff and Adrian del Caro (Cambridge: Cambridge University Press, 2001).

Human, All Too Human, ed. Richard Schacht, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1996).

Twilight of the Idols trans. R. J. Hollingdale (with *The Antichrist*) (Harmondsworth: Penguin, 1969).

Writings from the Late Notebooks, ed. Rüdiger Bittner, trans. Kate Sturge (Cambridge: Cambridge University Press, 2003).

The Will to Power, ed. Walter Kaufmann, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage Books, 1968).

OTHER WORKS CITED

Cox, Christoph (1999). *Nietzsche: Naturalism and Interpretation* (Berkeley: University of California Press).

Gardner, Sebastian (2008). 'Nietzsche, The Self, and the Disunity of Philosophical Reason' (in this volume).

Gemes, Ken (2009). 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual' (this volume).

Havas, Randall (2000). 'Nietzsche's Idealism'. *The Journal of Nietzsche Studies* 20, 90–9.

Janaway, Christopher (1989). *Self and World in Schopenhauer's Philosophy* (Oxford: Clarendon Press).

____ (2003). 'Nietzsche's Artistic Revaluation', in José Luis Bermúdez and Sebastian Gardner (eds.), *Art and Morality* (London: Routledge, 260–76).

____ (2006*a*). 'Naturalism and Genealogy', in Keith Ansell Pearson (ed.), *A Companion to Nietzsche* (Oxford: Blackwell), 337–52.

____ (2006*b*). 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual', *Proceedings of the Aristotelian Society*, suppl vol. 80, 339–57.

____ (2007*a*). 'Guilt, Bad Conscience, and Self-Punishment in Nietzsche's *Genealogy*', in Brian Leiter and Neil Sinhababu (eds.), *Nietzsche and Morality* (Oxford: Oxford University Press, 138–54).

____ (2007*b*). *Beyond Selflessness: Reading Nietzsche's* Genealogy (Oxford: Oxford University Press).

Kant, Immanuel (1998). *Critique of Pure Reason*, trans. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press).

Leiter, Brian (1994). 'Perspectivism in Nietzsche's *Genealogy of Morals*', in Richard Schacht (ed.), *Nietzsche, Genealogy, Morality* (Berkeley: University of California Press), 334–57.

Leiter, Brian (2002). *Nietzsche on Morality* (London: Routledge).

Owen, David, and Ridley, Aaron (2003). 'On Fate'. *International Studies in Philosophy* 35, 63–78.

Richardson, John (1996). *Nietzsche's System* (New York: Oxford University Press).

Ridley, Aaron (1998). *Nietzsche's Conscience* (Ithaca/London: Cornell University Press).

____ (2000). 'Ancillary Thoughts on an Ancillary Text', *The Journal of Nietzsche Studies* 20, 100–8.

Schopenhauer, Arthur (1969). *The World as Will and Representation*, trans. E. F. J. Payne, 2 vols. (New York: Dover).

4

How to Overcome Oneself: Nietzsche on Freedom

Robert Pippin

I

Does it matter to Nietzsche whether individuals are free in any of the manifold senses that have been so important throughout the philosophical tradition (self-knowledge, voluntarist "spontaneity," self-realization, autonomy, freedom from external constraint, morality, rational agency, authenticity, "non-alienated" identification with one's deeds, power to do what one desires)? And if persons are or can become in any of these or any other sense free, *how* important is it to him that all or some or a small few should be able to attain such a state or to exercise such a capacity?

Even given the great variety of interpretations of Nietzsche and the variety of positions attributed to him, the problem of freedom, whether as a metaphysical issue or as a possible human aspiration in any of the above senses, does not seem to be one of Nietzsche's central concerns. He often gives the impression that he thinks discussions of such topics are pointless and are motivated in self-deceit. However, it would appear that there is a sense of freedom, one at least somewhat still connected to much of our intuitive and everyday understanding of freedom, that is quite important to Nietzsche. It is the topic he discusses under the label: "self-overcoming." Or so I want to argue here.

But to get to that issue one should note first some of the details of his impatience with the "problem of free will" and with assessments of the "value of freedom," especially as those have been understood in Christian apologetics and in Western metaphysics from Augustine to Kant and Schopenhauer. Perhaps it would be better to say that his only interest in such questions is in dissolving the problems, not resolving them. In a much cited discussion of the metaphysical issue of the *causa sui* in section 21 of *Beyond Good and Evil*, he not only dismisses the voluntarist or incompatibilist commitment to such a *causa sui* (which he calls "a type of logical rape [*Notzucht*] and abomination [actually

just '*Unnatur*' in the original]''),[1] he goes on to be equally severe about what would appear to be its metaphysical contrary, the ''unfree will'' or the position of determinism. Both positions are said to be mere ''mythologies''; and he encourages us to see that the real problem (the problem as it is manifested in ''real life'') is the distinction and contestation between ''strong'' and ''weak'' wills. Even this will not be easy to understand since in BGE 19 Nietzsche had already effectively dismantled and rejected what would seem to be all the elements necessary for claiming that there is any faculty of the will, distinct from thought and desire. He reinterprets what had been taken as ''the will'' as in reality a ''complex of feeling and thinking'' which can produce a distinct sort of affect and pleasure (in commanding). It is this affect that is mistakenly interpreted as ''the will.''[2]

II

In passages like these, we can at least see what *does* interest Nietzsche about this and all other traditional philosophical positions: an etiology and often genealogy of the psychological type to whom one or the other of these positions would appeal. As in many other cases, philosophical positions are treated as psychological symptoms, and so an invitation to speculate on the need one type or another would have in order to believe either in a self-causing spontaneity or in what Nietzsche calls the ''dominant mechanistic stupidity'' of causal necessity. In this case he speculates on the stake a sort or type or ''race''—''the vain race''—would have in taking absolute credit for their deeds, and the stake in those motivated by self-contempt and so an interest in ''shifting the blame'' from their contemptible selves. They (the determinists) disguise this ''personal'' need when they write books, Nietzsche says, and adopt ''their most attractive disguise . . . socialist pity.''[3]

This same psychological inflection is apparent in his treatment of the political aspiration to achieve a free life, a life wherein one would be able to develop one's capacities and/or pursue one's preferences with minimal external constraint by others. In BGE 260, at the end of his concise summary of the difference between master and slave morality, he applies this psychological typology and suggests that a desire for freedom from external constraint is typical of the slavish, while by contrast ''artistry and enthusiasm in respect and devotion'' characterize the noble type.[4]

[1] BGE 21, p. 21. [2] BGE 19, p18–19; KSA, v. 31–4. Cf. also D 103.
[3] BGE 21, p. 22.
[4] Cf. Z 63: ''Free *from* what? As if that mattered to Zarathustra! But your eyes should tell me brightly: free *for* what?''

This psychological dimension—in effect the lens through which Nietzsche considers "the problem" of freedom, including the issue of "self-overcoming"—immediately introduces the most fundamental issue in interpreting Nietzsche. Although Nietzsche is capable of stating his basic interests and overall approach in a variety of not always consistent ways, there does indeed seem to be good evidence, like the passages just discussed, that his stated insistence on the priority of "psychology" in both explanation and evaluation is genuine, not another mere "mask." I mean such sweeping claims as in BGE 23, where Nietzsche encourages us to "clench our teeth," "open our eyes," and "keep our hand firm on the helm." We are to make a voyage that will entitle us to demand that

> . . . psychology be recognized again as the queen of the sciences, for whose service and preparation the other sciences exist. For psychology is now again the path to the fundamental problems.[5]

This does not appear to be an empirical psychology (Nietzsche presents the results of no studies), nor is it the kind of inference from a broad claim about "human nature" that we might expect from a philosophical anthropology. There are, though, some general characteristics of such a Nietzschean psychology with which most Nietzsche commentators would agree. First, he is primarily interested in what we need to say "psychologically" in order to understand what happens when we act on the basis of some value claim, or express in some way a commitment to a value. (One way of interpreting Nietzsche on the "priority" of psychology is already visible: he clearly believes that *any* activity, whether theoretical or practical, already involves such a value commitment (for example, to the "value of truth") and so the place of value and its psychological conditions in the economy of the soul must be "fundamental" for an account of any other activity).[6]

Second, it is often said that Nietzschean psychology must be "naturalist" and third, that it is therefore largely deflationary. At a minimum, the naturalism requirement amounts to an insistence that, when trying to account for the

[5] BGE 23.

[6] Some (and I think Nietzsche would be one of them, were he not to object so much to this abstract language) would claim that even the most basic form of intentionality requires the play of normative commitments. (To be conscious of X is to claim or judge that X, and so to undertake these commitments and be prepared to justify them if challenged.) It is in this sense that the primordiality issue in Nietzsche has to do with value, and in this case this means something like the basic authority of the constraints and requirements I undertake to accept, impose on myself and *hold to*. To assert, or to pledge to do, or to claim to know, are thus *au fond* kinds of promises, and such promises are not explicable as merely natural events. The constraints we undertake are not physical impossibilities. They require my futural commitments and my holding to them, sustaining them (under some understanding of, and commitment to, why I ought to), in order to be the promises—the assertions and expressions of intention—that they are. These sorts of commitments are thus basic or constitutive for the very possibility of thought, belief, action, all intentionality. And like many, Nietzsche would like to understand the source of this normative authority, why and in what sense we are bound as we are. Cf. GS 335.

human capacities required when persons direct their actions on the basis of norms, we should appeal to capacities also discoverable in non-moral or non-ethical contexts, and those capacities must be consistent with our being nothing but organic material bodies located in space and time. If we can only explain normative constraints and a set of practices by appealing to a capacity uniquely required by a particular view of value (such as a free will, an uncaused cause, or a unified subject independent of and directing its deeds), especially if that capacity is supernatural, the odds are high, at the very least, that we are dealing with a kind of philosophical fantasy.[7]

This enterprise turns out to be critical and deflationary, especially with regard to the set of values and practices that Nietzsche designates as "morality"—the Christian and post-Christian values of universal equality, absolute individual responsibility, and guilt. The way the psyche works in commitment to and pursuit of moral values is in reality[8] far different than the self-descriptions of moral agents. (*That* psychology is the sort of fantasy just described.)[9]

Finally though, if the Nietzschean enterprise is deflationary, it is not reductionist (or, we might say today, physicalist). One of the things natural organic beings can do, must do, is to create all sorts of different institutions under varying circumstances, train themselves to observe certain constraints and not others, and there is no reason to believe that exclusive attention to the biological or physical properties of these organisms best explains (or could explain at all) why they create one sort rather than another, and no way a purely natural science account could explain what these institutions actually mean to the participants, what they take themselves to be doing. (It *would* be right to say that *if* a reductionist naturalism were true, what it would be to be the subject of one's deeds would be to be, in effect, a spectator of the clashes and pulls and pushes of various contesting drives and passions.[10] But we cannot assume such a stance in leading or living a life; there is no first-person point of view which could embody such a perspective. We cannot "wait" to see which drive wins, but rather

[7] This is how Williams (1994) encourages us to understand what he calls Nietzsche's "minimalist psychology."

[8] I mean in *historical* reality. The "real" roles of *ressentiment*, hatred, and revenge are not instances of any general law about the psyche, but aspects of a "slavish" institution at a time. To deny that this restriction to historical time could be a possible explanation (because linked to something historically unique) is like saying we cannot discuss why Emma Bovary had an affair with Leon unless we see it as an instance of a general law about bored provincial housewives.

[9] This is a minimalist ascription of naturalism to Nietzsche, but it is, I think the only sort justified. The topic is a complex one, though, and Nietzsche has different things to say about it in different contexts. For example, the famous "translate humanity back into nature passage" in BGE 230, and yet the attack on the Stoics, on the idea that one could actually live "according to nature," (given that nature is mere "indifference") in BGE 9. I agree with many of Christopher Janaway's objections to more than minimal naturalist reading of Nietzsche in his 'Naturalism and Genealogy' (2006).

[10] See the description in Velleman (1992: 461) and the discussion of this and the general problem of "fatalism" in Nietzsche in Leiter (2001: 221–54).

must act, and Nietzsche has both already dismissed the determinist version of a naturalist position, and in II 12 of *On the Genealogy of Morality* had insisted that a "*Grundbegriff*" for his own notion of "life" is "*Aktivität*.")[11]

III

These are all quite contestable claims, and the status of a Nietzschean psychology is easily a book-length topic.[12] But since *so* much depends on how Nietzsche understands a psychological treatment of the nature and value of freedom, we should also keep in mind that Nietzsche himself gives us a model to think of when we are interested in his notion of psychology, and this model will be important in assessing how he wants us to think about the issue. I mean the sixteenth- and seventeenth-century French moral psychologists, whom Nietzsche cited with such great enthusiasm, especially and above all, Montaigne (about whom he had almost nothing critical to say in the forty-eight references to him), but also, in a more qualified, yet still enthusiastic way, Pascal and La Rochefoucauld. (Pascal's "*l'homme honnète*" is the clear model for the Nietzschean "*freie Geist*" from the 1876 *Nachlass* written in preparation for *Human All Too Human*, and thereafter.[13] I think that there are two things that interest Nietzsche about this model.

What intrigues Nietzsche in "essays" or "maxims" or "pensées" is that they are presented without, and with no hidden reliance on, any putative deeper philosophical theory of being, or of human nature or of reason or of anything else, and it is clearly an assumption in all three (and by Nietzsche) that this is not a limitation undertaken out of modesty but unavoidable if one is to write "honestly" about human beings as they are, as they live, and so this assumed priority of a psychological perspective counts as a virtue. (The BGE passage had claimed that psychology was to be *first* philosophy, not an enterprise that rested on metaphysics or a substantive theory of nature.) And it is also no accident that the three moralists Nietzsche admired the most wrote in such unusual, original forms, quite foreign to any notion of systematic thought or metaphysical or epistemological foundations.

Not that it is easy to imagine what psychology as first philosophy might look like, how the issue of what is at stake psychologically in posing a question or holding a belief could be understood such that it would *not* itself have to rest

[11] GM p. 52. (Not that it is entirely clear what Nietzsche takes "*Aktivität*" to commit him to.)

[12] See Pippin (2006a). The following discussion of the French is a summary of some of the themes of this book.

[13] Vivetta Vivarelli (1994) has, I think, established this convincingly in 'Montaigne und der freie Geist.' Cf. also *Human All Too Human* 50 for Nietzsche and La Rochefoucauld on pity. For criticisms of La Rochefoucauld as still bound to essentially Christian categories of evaluation, see in the *Nachlass*, (1880–1), KSA, v. ix., p. 295 and (1881–2) KSA, v. ix, p. 441.

on some view of being or nature or mind and so forth, some view of how what Nietzsche claims about psychology is *possible*. Even if Nietzsche does not mean *that* very strong or logical notion of priority, and means only to signal something weaker—say, that he thinks these psychological questions are just more interesting, have too long been neglected and have led to self-deceit, or are in some way more important, and so forth—we will still need to hear about this scale of interest or importance. All of this remains mostly implicit. But he at least issues a number of clear warnings about what will happen if we get his focus wrong. The most interesting is in Part Two of HAH, in the 'Assorted Opinions and Maxims' and describes what Nietzsche calls "An original sin (*Erbsünde*) of philosophers." It is important enough to quote at length.

> Philosophers have at all times appropriated the propositions of the examiners of men (moralists) (*Menschenprüfer*) and *ruined* them, inasmuch as they have taken them for *unqualified* propositions and sought to demonstrate the absolute validity of what these moralists intended merely as approximate signposts or even as no more than truths possessing tenancy only for a decade—and through doing so thought to elevate themselves above the latter . . . . Even the word 'will', which Schopenhauer remoulded as a common designation for many different human states and inserted into a gap in the language—greatly to his advantage insofar as he was a moralist, since he was now at liberty to speak of the 'will' as Pascal had spoken of it—even Schopenhauer's 'will' has, in the hands of its originator through the philosopher's rage for generalization turned out to be a disaster for science . . . finally so that it (Schopenhauer's 'will') can be pressed into the service of all kinds of mystical mischief it has been misemployed towards a false reification . . . .[14]

Secondly, what Nietzsche seemed to get from this reading of *les moralistes*, especially in the late 1870s, is a way to formulate one of his most important questions: while, according to Nietzsche, La Rochefoucauld's tendency to see petty egoism everywhere, while a mark of great honesty, finally belittles man unfairly, and while Pascal's noble and honest and eventually despairing soul was crushed by Christianity or the Christian understanding of the weakness and depravity of man, by contrast the Nietzschean question is at its clearest with Montaigne. *How*, he wants to know above all, did Montaigne manage to exhibit such a thorough-going skepticism and clarity about human frailty and failings (the virtue of *Redlichkeit* so often praised by Nietzsche) *without* Pascal's despair and eventual surrender[15], or La Rochefoucauld's icy contempt for the "human all too human"?[16] Instead Montaigne ended up a thoughtful, ferociously honest, cheerful free spirit, someone who had succeeded at what seems a supreme

[14] HAH 215–16.

[15] BGE 62 enlists Pascal in Nietzsche's own view of Christian degeneracy; and there are pensées that could have been written by Nietzsche. See the famous "Le moi est haïssable" paragraph; §597, and Pascal's remarks about the two qualities of the self (Pascal: 2000, ii. 584).

[16] Cf. HAH 36.

Nietzschean goal, the task, "to make oneself at home on the earth" ("... sich auf der Erde heimisch zu machen.").[17]

So what would avoiding such a "false reification" and achieving something in the spirit of the French moralists look like? It is at least clear that Nietzsche believes that views of the soul and its capacities vary with views about norms; normative commitments are subject to radical historical change and vary among different types in any epoch; and so what counts as soul or psyche or mind also changes. (So no reification; moralists deal in truths that are time-indexed, holding for perhaps only "a decade.") "The soul" is merely the name for a collective historical achievement of one sort or another, what we have made ourselves into at one point or other in the service of some ideal or other. When we describe to each other what we think the soul is, we express thereby a view about determinate ideals and their psychic functioning; we often are thereby oriented in our explanations from something like psychic health. Hence also the deep interconnection or inseparability between psychology and (an evaluative often deflationary) genealogy.[18]

IV

But all this is by way of setting the context, of beginning to understand how Nietzsche wants to raise the question of freedom. Let us say that he seems primarily interested in freedom as a value or aspiration, has his own views about the nature of genuine freedom, and he is especially interested in the psychology and psychological typology that would help explain genuine freedom (the "psychology of freedom" one might say), and the psychology behind differing understandings of freedom, what is at stake in the appeal of one or another aspiration, all presumably with the hope that we might eventually see what sorts of ideals and aspirations might be admirable and which not (although again the connection between diagnosis and evaluation, here as elsewhere in Nietzsche, is not obvious). Finally, whatever this psychological treatment is, it is

[17] 'Schopenhauer as Educator,' in UM p. 135. KSA, i. 348. See also HAH 36 on Le Rochefoucauld and the "danger" that Montaigne avoided.). Also, D 31, p. 46. *Doubt about doubt.*—'What a good pillow doubt is for a well-constructed head!'—This saying of Montaigne's always provoked Pascal, for no one longed for a good pillow as much as he did. Whatever was wrong?

[18] Cf. GS 335 for one of Nietzsche's most interesting reflections on the relation between a claim like "this is right" and the question "what impels me to listen to it?" As Zarathustra says, living is leading or directing a life and so living essentially *is* esteeming (*schätzen*) or valuing. So the priority of psychology looks very like Nietzsche's occasional suggestion that all human phenomena, including knowledge claims, be looked at "from the perspective of life," that is, as valuing. Psychology, that is, will make manifest what is involved, what we are committed to, in understanding "living" in this sense. As has already been noted, from the time of the second *Untimely Meditation* on, Nietzsche claimed that modern philosophy and science and religion and even art had lost any connection with life in this sense, and insofar as any of these enterprises had a living presence in modern culture, they were actively life-denying presences.

neither an empirical psychology, nor a philosophical psychology modeled either after Aristotle or Hume, and seems modeled after the older French sense of a *moraliste*.

Some aspects of what results from this approach to freedom are of course well known. If herd morality, conformism and sheep-like timidity are to be held in contempt, then some contrary notion seems suggested, some ideal of *social independence* and a kind of *self-rule* or *self-reliance*. A Stoic like emphasis on self-rule is particularly prominent in later works, *Twilight of the Idols* especially. There in section 38 of the 'Expeditions of an untimely man' (*Streitzüge eines Unzeitgemässen*) section, called 'My conception of freedom,' Nietzsche offers an encomium to a wide-ranging set of psychological dimensions of freedom.

> That one has the will to self-responsibility. That one preserves the distance that divides us. That one has become more indifferent to hardship, toil, privation, even to life. The man who has become free . . . spurns the contemptible sort of well-being dreamed of by shopkeepers, Christians, cows, women, Englishmen and other democrats. The free man is a warrior.[19]

The passage goes on like this, praising danger, risk and strength, but, as he tries to characterize what he calls "psychologically true" [*psychologisch wahr*] about freedom, Nietzsche adds something that is easy to overlook.

> How is freedom measured, individuals as in nations? By the resistance which has to be overcome, by the effort it costs to stay aloft. One would have to seek the highest type of free man where the greatest resistance is *constantly being overcome*.[20]

Nietzsche here is most interested in a sort of psychological self-relation as constitutive of freedom, (a self-relation that would immediately define of itself acceptable and unacceptable relations to others) and he clearly thinks of this psychological state as an achievement along a spectrum of possibilities, not an *either-one-has-it-or-one-doesn't* kind of capacity, as among the voluntarists, an achievement that he also treats in a sort of soul-writ-large way, ascribing this achievable state to possible nations as well.[21] And he notes that whatever the resistance that has to be overcome, there results no settled state, the resistance must be constantly (*beständig*) overcome (*überwunden*). (He notes that he understands freedom as the Romans and Venetians did, as "something one has and does not have, something one wants, something one conquers . . ."[22]) Earlier in TI,

[19] TI p. 92. [20] Ibid, my emphasis.

[21] On this account, freedom for Nietzsche involves achieving a certain sort of relation to oneself that only a few are capable of, and, it would appear, which cannot be the product of any sort of direct intention by an individual. The conditions for its achievement depend on far more than individual "will power." This means such a notion of freedom does not rely at all on a libertarian or "could have unconditionally done otherwise" conception of freedom. See the valuable discussion by Ken Gemes (2009), especially his case for this "achievement" sense of freedom and Nietzsche's compatibilism. See also my discussion in Pippin (2006a).

[22] TI p. 93.

in section 3 of 'Morality as Anti-Nature,' he had stressed both this notion of achievement and also characterized it as an odd sort of calm amidst unsettled endeavors.

> Or the expression of ripeness and mastery in the midst of action, creation, endeavor, volition and a quiet breathing, 'freedom of the will' attained [*die erreichte Freiheit des Willens*] . . . *Götzendämmerung*: who knows? Perhaps this too is only a kind of 'peace of soul' [*Frieden der Seele*] . . .[23]

V

But what *is* this sort of self-relation? What counts as self-mastery in this sense? Two elements have been suggested in recent discussions. One is inspired by passages in GS, echoed elsewhere, about what appears to be a kind of literary self-creation as the "greatest will to power," the desire "to impose upon becoming the character of being."[24] On this view freedom for Nietzsche has elements both of a self-realization theory ("become *what* you are," the subtitle of *Ecce Homo*, i.e., become, own up to being, the kind of creature who must fashion his own character and personality) and an authenticity theory (as in passages in 'Schopenhauer as Educator' and GS which insist that one ought to become who one individually is—"be yourself" (*sei du selbst*) in the former, and "You should become who you are" (*Du sollst der werden, der du bist*) in the latter.)[25] And according to these interpretations, these injunctions can be fulfilled if one creates oneself as if a "literary character" in a novel, "gives style" to one's character, finds a way to identify oneself with all one's actions, to see that no aspect of one's character and deeds can be what it is without literally every other aspect of one's character and deeds. This would be one way of attaining what appears to be the second necessary aspect of "*erreichte Freiheit*," "achieved freedom," a complete and hierarchical *unity* among states of one's soul, memories, desires, aversions, and so forth. We would thus have fulfilled what appears to be the ideal suggested by GS to become "the poets of our own lives," or the call to "become those who we are," human beings who, Nietzsche says, "create themselves."[26] We would not be what circumstances or others have made us, not be pulled or pushed about by whatever inclination or aversion we happen to be feeling, would be who we really, ontologically and individually, are, self-creating and individually self-created, hierarchically unified beings, and in that sense would have "reached" freedom.

There are several problems with this approach. In the first place, the formulation attributes to us a distinct power or capacity which seems to cry out for a further, deeper metaphysical account, something, I have suggested, inconsistent

[23] TI p. 44. [24] WP 617. [25] UM p.127; GS p. 219. [26] GS 299, p. 335.

with Nietzsche's claim about the priority of psychology and too traditional to match his intentions. Why are we entitled to believe that we *can* assume such a possible independence between a creator or ruler self and a created object and "commonwealth"?[27] It seems just as implausible that one *could* assume such a detached, artist-like creator position as that, for any deed, one could have done otherwise.

Secondly, some aspects of the position seem wildly implausible. *Why* is it a condition of this literary unity that I must be able to understand *every* single deed of mine as necessary for another? That my getting a divorce or resigning my professorship is to be somehow understood as linked together with which tie I choose to wear or what I had for breakfast?

Third, the self-creation view often imports a notion of what *creation* is that is foreign to Nietzsche, whose sympathies here are not with the hyper-modernist, ironic, insubstantial position of Rameau's nephew or Proust, but remain essentially romantic. From the beginning of his publishing career to near the end, the creative state is always understood as Dionysean, a dissolution of boundaries and not their Apollonian establishment, a state of reverie and intoxication (*Rausch*).[28] In the 'Zarathustra' chapter of EH he goes to great lengths to describe precisely the involuntariness and necessity of this creative state, insisting that he, as author, has virtually *vanished*,[29] and his account of artists in BGE, those for whom necessity and "freedom of the will" (cited with the usual sneer quotes by Nietzsche) are the same, makes the same point even more clearly.[30] Even in the GS 335 self-creation passage, Nietzsche goes on to say that such creation requires that we learn "everything that is lawful and necessary in the world," that to become creators we have to become physicists.[31] (In the discussion of the "way of the creator" in *Zarathustra*, Nietzsche's account does not sound like aesthetic self creation: "With my tears go into your loneliness, my brother. I love him who wants to create over and beyond himself and thus perishes."[32]

Finally, and most important, the concept of creating oneself as if the author of one's life is inherently unstable, potentially incoherent. In literary terms, the character creating the unity of character in the story of one's life is obviously also a character himself *in* the story he is narrating. For *that* character to form a unity with the character being created, one will always require, in an obvious iteration, another creator-character who could bring the creator and created characters together, and so on. There are literary attempts to write the story of a character who finally becomes the author of the story of this becoming that

[27] Cf. BGE 225. "In human beings creature and creator are combined . . .", p. 117.

[28] TI p. 71. Nietzsche can speak out of both sides of his mouth on this, praising *both* irony and whole-heartedness. Cf. BGE 40, 284, 289, and contrast TI 44. As we shall see, the tension created by such passages is not an oversight by Nietzsche. He very deliberately wants to claim both aspects as essential in "the achievement of freedom."

[29] EH p. 300. [30] BGE 213. [31] GS p. 266. [32] Z p. 65.

we have just read. Proust is the obvious example. But the mere existence of that novel simply raises its own version of this problem, a legendary one for readers: how to understand the relation between the Marcel who is the object of the story, the older Marcel who appears to be narrating and writing the story, and the absent Ur-narrator, Proust himself. If anything, the briefest contemplation of the details of this issue in Proust make much more unlikely the possibility of construing freedom as the self-creation of a unified character. And all of this is not yet to mention that this account leaves unclear how we are to put together the idea of an author/creator of one's own character with the requirement of some whole-hearted identification with one's creation, something far more passionate and unqualified than any picture of a *Rameau*-like independence.

VI

We move in a more promising direction, I want now to suggest, if we pay attention to another dimension central to Nietzsche's picture of agency, something already alluded to in several passages cited above. That is the fact that for Nietzsche one does not count as an agent, the true subject of one's deeds, just by in effect "showing up." One has to *achieve* something—and I am suggesting that this is a distinct sort of psychological self-relation, both attitudinal and dispositional—in order to be capable of any *real* practical intentionality or real agency. (I say here "one has to achieve . . ." but as we shall see shortly it is important to note that there is no reason to think that Nietzsche must mean an individual achievement, the *result* of an individual's "resolve" and efforts. It could be, perhaps exclusively, a civilizational or social achievement, or even an "achievement" of fate that is responsible for one's being in such a self-relational state.) The state is described in any number of ways, but all them have something to do with a kind of self-dissatisfaction, a generally "negative" as well as positive stance towards some current set of standing attitudes, commitments, and ideals. By and large, Nietzsche describes this condition when he wants to talk about, as he calls it, freedom of the *highest* sort, the *true* or paradigmatic instance of independence from others and a kind of self-direction, not humdrum or ordinary cases of such intentionality, like pumping water, or turning on the light switch, and so forth. All action involves a negation of a sort, an alteration of what would have remained the same without one's intervention, but Nietzsche appears particularly interested in a kind of inward-looking self-negation, a transformation of what had been the subject's restraints, or commitments, basic desires or passions, all in a way that makes possible a new kind of outward-looking relation to the world. In those paradigmatic cases (where, especially the direction and course of one's whole life is at stake) he often focuses our attention on what he calls a "tension of the spirit" that allows a genuine "self-overcoming."

One initial, still quite crude summary of what he is getting at in these passages would simply be that achieved freedom involves achieving a capacity both to sustain a whole-hearted commitment to an ideal (an ideal worth sacrificing for, that provides the basis for a certain hierarchical unity among one's interests and passions), and what appears at first glance to be a capacity in some tension with such whole-heartedness—a willingness to overcome or abandon such a commitment in altered circumstances or as a result of some development. To be unable to endure the irresolvable dynamic of what Nietzsche calls an ideal's or a goal's or a value's constant self-overcoming, to remain dogmatically attached to an already overcome form of life (as with the Christians described in *The Anti-Christ* who, as a result of the "self-overcoming of the intellect" know that prior terms of belief can no longer be invoked in the same way but do so anyway, "Everyone knows this: *and everyone nonetheless remains unchanged*")[33]. Or to concede such mutability but with a cynical relativism that prohibits any wholehearted identification with a new ideal (as the "last men," like those who so casually respond to the "God is dead" news in GS 125), or to slide into a complacent, lazy identification with whatever is conventionally valued (as Nietzsche says in the second *Untimely Meditation* of the Germans of his own day, who have let themselves go and "elect for ease and comfort and the smallest possible degree of self-overcoming"),[34] all these are treated as forms of unfreedom. To be sure, freedom is not the term Nietzsche prefers, although as we have seen, despite its dangerous associations, it is one he uses.[35] He more often speaks of satisfying the conditions of life, leading a life, truly living, recognizing one's life as one's own. The Nietzschean "theory of agency" is couched in such elusive formulations as: "I wish your self were in the deed like the mother is in the child; let that be your word on virtue" from *Zarathustra*.[36] But the agency issues seem to me clearly present. This is so even though neither Nietzsche nor Zarathustra ever simply encourages us to "overcome yourselves." (The issue seems to be the proper acknowledgement and endurance of the self-overcoming character of life, an orientation that itself, as we shall see, has several social and historical conditions for its possibility.)

The achieved state of mind that Nietzsche promotes in these passages is not easy to make out. In the first place, underlying it appears to be a much broader theory about the historical fragility of all human norms, the inevitability not just of a kind of organic growth and death, but of a self-undermining process that sometimes sounds positively Hegelian. It is this historical fate for norms that *requires* the kind of acknowledgement and endurance that Nietzsche praises

[33] A 38, p. 150. [34] UM, p. 80.

[35] The well-known discussion of the "sovereign individual" in the second essay of GM II is another case in point, as is his discussion of the "right" to make promises (GM II 1). See my discussion of the general issue raised by such claims about promising in the second chapter of my *Nietzsche, moraliste français*.

[36] Z p. 96.

when he discusses self-overcoming. Indeed some of the references translated as self-overcoming are actually to the famously Hegelian term "*Selbstaufhebung*" and its cognates, as in GM when Nietzsche claims that "every good thing on earth" eventually overcomes ["sublates"] itself, or later in GM when Nietzsche proclaims what he calls the "law of life," "the law of the necessity of self-overcoming in the nature of life" where he uses both *Selbstaufhebung* and *Selbstüberwindung*.[37] Morality's commitment to an ethic of truthfulness about intentions is his chief example of this self-undermining and self-overcoming, but as stated it is presented as simply a law of life itself. (The death of Attic tragedy, which Nietzsche calls a suicide, a *Selbstmord*, might be another example of this law. So might truthfulness about the value of truth.)[38]

Secondly, the state itself, the proper responsiveness to the self-overcoming character of life, is itself quite complicated, full of dialectical, affirmation/negation flourishes, all of which evoke Nietzsche's characterization of freedom in TI 38, that it is something one "has and has not." When Zarathustra discusses the "way of the lover," he characterizes it as "yourself you love and therefore you despise yourself, as only lovers despise. What does he know of love who does not have to despise precisely what he loved."[39] As in HAH, Nietzsche will frequently pronounce himself in favor of a morality "as a continual self-command and self-overcoming practiced in great things and in smallest,"[40] but his more detailed accounts of such a self-relational, self-overcoming state are even more figurative and more difficult to summarize. A typical passage is from his account of *Zarathustra* in EH.

> The psychological problem in the type of Zarathustra is how he that says No and *does* No to an unheard of degree, to everything to which one has so far said Yes, can nevertheless be the opposite of a No-saying spirit . . . how he that has the hardest, most terrible insight into reality, that has thought the 'most abysmal idea', nevertheless does not consider it an objection to existence, not even to its eternal recurrence—but rather one reason more for being himself the eternal yes to all things . . .[41]

The language of a negation that is also an affirmation is sometimes transposed into a more familiar Nietzschean trope and one more familiar in compatibilist accounts of freedom, the simultaneity, in the experience of true freedom, of both freedom (the capacity to negate, free oneself from, some state or other)

[37] GM III, p. 27.

[38] In Hegel there is an elaborate speculative theory about why commitment to a certain norm or to any principle would, in the process of becoming a more articulated or self-conscious commitment, render continued allegiance more and more problematic until the original position was "negated" even while preserved and "raised up," or *aufgehoben*. There is no such theory in Nietzsche although there are indications that he thinks that commitments require a certain sort of selective attentiveness, perhaps even a sort of self-delusion which is inevitably rendered more difficult with time. Cf. his summary of this "process" in EH: "The self-overcoming of morality, out of truthfulness; the self-overcoming of the moralist into his opposite—into me—that is what the name of Zarathustra means in my mouth," p. 328.

[39] Z p. 65. [40] HAH p. 322. [41] EH p. 306.

and necessity (the affirmation of a cycle of necessity). In the 'Richard Wagner in Bayreuth' section of UM, Nietzsche claims that in listening to Wagner, one does not have the usual aesthetic experiences of enjoyment or interest,

> . . . only feels only the necessity of it all. What severity and uniformity of purpose he imposed upon his will, what self-overcoming the artist had need of in the years of his development so as at last in his maturity to do with joyful freedom what was necessary at every moment of creation, no one will ever be able to calculate: it is enough if we sense in individual cases how, with a certain cruelty of decision, his music subordinates itself to the course of the drama, which is as inexorable as fate, while the fiery soul of this art thirsts to roam about for once unchecked in the freedom of the wilderness.[42]

The two sorts of formulations, necessity and freedom, affirmation and negation, are brought together in a still highly figurative passage in GS 276.

> . . . I, too, shall say what it is that I wish from myself today, and what was the first thought to run across my heart this year—what thought shall be for me the reason, warranty and sweetness of my life henceforth. I want to learn more and more to see as beautiful what is necessary in things; then I shall be one of those who makes things beautiful. *Amor fati*: let that be my love henceforth. I do not want to wage war against what is ugly. I do not want to accuse; I do not even want to accuse those who accuse. Looking away shall be my only negation. And all in all, and on the whole: some day I wish to be only a yes-sayer.[43]

Finally, there is a last set of images that appear to treat as a condition for such a self-overcoming, that is, self-negating and yet self-identifying and self-affirming, state. These images continue the unusual emphasis of the passage just cited, where the achievement of freedom seems much more the achievement of an intellectual (i.e. "to learn more and more") and erotic attitude (as in the references to his heart and what he hopes for his "love"). This is consistent with the intellectualist account of freedom in Socraticism, Stoicism, and Spinoza, a notion which Nietzsche expressed admiration for. For Nietzsche too there *is* a kind of knowledge that will set one free, but it is not knowledge of the human good and not, or at least not wholly, the Spinozist knowledge of necessity. It appears to be a psychological realization of the ineliminable need for self-overcoming. The image is also consistent with the fact that, despite what can seem the hortatory character of Nietzsche's rhetoric, many of the passages we have looked at do not really directly encourage readers to *do* anything, as if simply to resolve to become free, to attain freedom. One cannot, as an act of will traditionally understood, will oneself into a state of knowledge or to desire something. The conditions for the attainment of freedom—the proper relation of attachment and detachment—seem largely prevoluntary and extend in scope beyond what individuals can do. Likewise, while Nietzsche is not encouraging

[42] UM p. 244. It is interesting that, when describing his own most difficult "self-overcoming," his break with Wagner, Nietzsche describes it both as a difficult act of self-overcoming, but as also his fate (*Schicksal*). *The Case of Wagner*, p. 155.

[43] GS p. 223.

anyone to "overcome himself," but rather writes abut bearing or enduring a fate in a certain way, he is *also* still not encouraging one not to flee that fate, as if that too were a matter of simple resolve, but describing what it would like for that fate to be borne or endured and affirmed, what else would have to be "in place" for that to happen, and what it would be like.[44] Thus the following images that suggest these necessary conditions.

For example, in the Preface to BGE, he notes that our long struggle with and often opposition to and dissatisfaction with our own moral tradition, European Christianity, has created a "magnificent tension (*Spannung*) of the spirit in Europe, the likes of which the earth has never known: with such a tension in our bow we can now shoot at the furthest goals." But, he goes on, the "democratic Enlightenment" also sought to "unbend" such a bow, "make sure that spirit does not experience itself so readily as 'need.' "[45] This latter formulation coincides with a lapidary expression in GS. In discussing "the millions of Europeans who cannot endure their boredom and themselves," he notes that they would even welcome "a craving to suffer" and so "to find in their suffering a probable reason for action, for deeds." In sum: "neediness is needed!" (*Not ist nötig*.)[46] In Z, the point is formulated in a similar way.

> Beware! The time approaches when humans no longer launch the arrow of their longing beyond the human, and the string of their bow will have forgotten how to whir! . .
>
> Beware! The time approaches when human beings will no longer give birth to a dancing star. Beware! The time of the most contemptible human is coming, the one who can no longer have contempt for himself.[47]

There is no suggestion in any of these passages that "neediness" can be created by an act of individual will and the conditions described as necessary for self-dissatisfaction and so self-overcoming (in Nietzsche's somewhat purple prose, "self-contempt" and "self-despising") are clearly here historical (dependent on one's time) and social (dependent in some way on the state of a shared social world).

There is a kind of culmination of this sort of language in the section, 'On Self-Overcoming' in Z. In this section we hear again many of the themes sounded about self-overcoming. "Life" reveals to Zarathustra its "secret"—"I am that which must always overcome itself."[48] Any good and evil presumed not transitory

[44] This dimension obviously has many resonances in the Heidegger of *Sein und Zeit*. Two are the clearest: Heidegger's description of authenticity as a "readiness" for anxiety, and so the stance of "anticipatory resoluteness" for a being that is always "ahead of itself"; and the fact that Heidegger does not think of the forgetfulness of the everyday as a "failure" of Dasein that we should be encouraged to avoid. Rather such forgetfulness is as constitutive of what it is to be Dasein as authenticity. There is something of this in Nietzsche's discussion of the "clouds of illusion": necessary for the creation and affirmation in life.

[45] BGE p. 4. [46] GS 56, p. 117.

[47] Z p. 17, (translation altered). See also 'On Unwilling Bliss' in the third part, where Zarathustra speaks of the "desire for love."

[48] Z p. 115.

are said "not to exist." "Driven on by themselves, they must overcome themselves again and again."[49] But now something else, apparently momentous, is added. We also learn that it is *this* feature of life—and Nietzsche seems to mean here this feature of the historical life of values, the feature of having to overcome itself—that is somehow equivalent to the claim that "all life is will to power."[50]

We have heard this link before, as in the HAH remark about "continual self-*command and* self-*overcoming* practiced in great things and in smallest." But the conjunction of topics is puzzling. One set of images deals with the necessity of mastery and servitude in existence, the omnipresence of commanders and obeyers and such maxims as "Whatever lives, obeys," and ". . . he who cannot obey himself is commanded." The second deals with images of the transitoriness of any fixed, settled value and suggests the great difficulty of acknowledging, accepting, or "living out" in some way this perpetually self-undermining dynamic. Both, this unavoidable struggle for mastery and this ability to acknowledge the transitoriness of that in the name of which one claims mastery, seem related to Nietzsche's view of freedom.

The link between the will to power and self-overcoming appears to be related to the unusual way Nietzsche understands power, or more precisely what he is willing to count as the realization of any will to power. For Nietzsche sometimes concedes that the most essential element in a contestation over power has to be the interpretive question of what *counts* as having achieved mastery. (". . . all events in the organic world are a subduing, a becoming master, and all subduing and becoming master involves a fresh interpretation, an adaptation through which any previous 'meaning' and 'purpose' are necessarily obscured or even obliterated.")[51] The ability to bully and tyrannize someone into cooperation is one thing, the ability to inspire true service another; self-command is one thing, self-overcoming is another; being unimpeded in the satisfaction of one's desires is one thing; being able to order one's desires in a "hierarchy of rank" is another; commanding is one thing, being "strong" enough to "yield" command is another.

"Yield" (*hingeben*) is Nietzsche's word in the passage that link the themes of mastery and self-overcoming.

> And as the smaller yields to the greater that it may have pleasure and power over the smallest, thus even the greatest still yields, and for the sake of power risks life. That is the yielding (*Hingebung*) of the greatest; it is hazard and danger and casting dice for death.[52]

The upshot of these obscure allusions seems to be that *nothing* really counts in some probative way as "*the*" establishment of mastery. There are of course wider, more apparently metaphysical dimensions of the will to power notion in

[49] Z p. 116.

[50] I discuss the place and function of this ideal in Z as a whole in my 'Introduction' to the new Adrian del Caro translation of Z (Cambridge: Cambridge University Press, 2006.)

[51] GM II 12, p. 77. [52] Z p. 115.

Nietzsche's work, but in the most intuitively obvious instance—dominating and being dominated in the human sphere—these passages suggest that it is a mistake to understand the meaning of mastery in such a human dimension without taking into account the unsettled and precarious interpretation of mastery upon which the claim of having-mastered will always actually rest. These "interpretations" are mutable, of a time and place, and so true masters must be prepared to "yield" as well as to seize "command." To be capable of this is to have achieved freedom; to become, in Nietzsche's earlier invocation of freedom, a *freier Geist*, a free spirit.[53] And all of this is apparently connected with the uniquely historical situation in which Nietzsche believes we must evaluate and act, the first epoch in which we must admit that we do not know, in the traditional objectivist or religious sense, what is worth wanting or aspiring to, where the danger of "nihilism" (on this reading, not a failure of knowledge or of will, but of desire) is always threatening. The prospect of a constantly "self-overcoming" structure of valuation is what obviously provokes this danger, and Nietzsche's aspiration is that such an age might also allow human beings who are prepared to be *constantly* "over" or beyond themselves—"*Übermenschen*," they might be called. He sums up such a stance in just this way in GS 347.

> . . . one could conceive of such a pleasure and power of determination, such a freedom of the will that the spirit would take leave of every faith and every wish for certainty, being practiced in maintaining himself on insubstantial ropes and possibilities and dancing even near abysses. Such a spirit would be the free spirit par excellence.[54]

This is a tangle of themes, any one of which would require much more discussion. I have suggested that we should follow Nietzsche's lead in considering the "problem of freedom" to be a "psychological" problem in his sense of the term. That is, Nietzsche clearly considers freedom to consist in some sort of affirmative psychological relation to one's own deeds, a relation of identification, finding oneself in one's deeds, experiencing them as genuinely one's own. He also considers this state of being an achievement, rather than the exercise of an inherent capacity. The achieved state in question requires an unusual intentional self-relation, in particular an intentional relation to one's own commitments. The relation involves both a kind of whole-hearted identification and affirmation, as well as the potential for great self-dissatisfaction. It is a state of great "tension." One is neither as passionately identified with one's projects as Goethe's Werther, nor as ironically detached from them as Diderot's Rameau's nephew. To be in such a state of tension is to be capable of self-overcoming, genuine freedom, or

[53] Cf. the discussion in GM II 10 where, in his genealogy of punishment, Nietzsche notes that "as its power increases, a community ceases to take the individual transgressions so seriously . . ." and that in the development of justice it all ". . . ends by winking and letting those incapable of discharging their debt go free; it ends as does every good thing on earth by overcoming itself," pp. 72–3.

[54] GS pp. 289–90.

one might better say, to be capable of bearing the burden of such self-overcoming and of affirming under its condition. Yet the conditions for the possibility of such an achievement extend far beyond what an individual alone can be called on to achieve. These conditions are partly social and historical, and Nietzsche's basic psychology does not appear to include an addressee for any call to simple resolve or strength of will. (Even if it did, a will cannot resolve to be strong unless it already is.) Whatever the nature of the perspective achieved by Montaigne that Nietzsche so much admired, it was not achieved (cannot be achieved) by force of will or as the consequence of practical reasoning. All of this is in partial explanation of Nietzsche's unusual rhetoric, a mixture of so many different styles, tropes, and voices, as he tries, in effect, to create a picture that can get some grip or hold on his readers. This is especially true of Zarathustra who, readers inclined to the blood-thirsty-blonde-beast Nietzsche rarely notice, is quite an unusual hero; often perplexed, confused, disappointed in his comrades and followers, at times bombastic and self-important, at times tender-hearted and resigned to his fate; a figure as much of parody as of tragedy, a fact and paradox Nietzsche was himself eager to point out. All of which raises the question that, in Nietzsche's terms, arises within the perspective of "life"—what would it be to live out both bearing the burden of the finitude and temporal fragility of one's ideals and yet to be capable of "self-overcoming"?

These questions can be pursued in any number of ways. My point here has only been to suggest if the topic is freedom and self-rule according to Nietzsche, any investigation of that issue will have to lead to yet another unusual paradox in Nietzsche's account. There is, in other words, an analogue here to the famous maxim of Pascal's that Nietzsche must have relished: "*La vraie philosophie se moque de la philosophie*." True philosophy ridicules, has nothing to do with, "philosophy." The Nietzschean turn of that screw would be: the true realization of the will to power has nothing to do with gaining and holding power as traditionally understood, except as an indifference to power in this sense. The implication from the passages we have reviewed would seem to be: the true realization of the will to power, genuine freedom, has rather to do with self-overcoming.

REFERENCES

WRITINGS BY NIETZSCHE

Kritische Studienausgabe, ed. Giorgi Colli and Mazzino Montinari (Berlin/New York: de Gruyter, 1980).

The Anti-Christ, trans. R. J. Hollingdale (Baltimore: Penguin, 1968).

Beyond Good and Evil, eds. R-P. Horstmann and J. Norman, trans. Judith Norman (Cambridge: Cambridge University Press, 2002).

The Case of Wagner, trans. Walter Kaufmann (New York: Vintage, 1967).

Daybreak, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1982).
Ecce Homo, in Nietzsche, F., *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmann (New York: Vintage, 1989).
On the Genealogy of Morality, trans. Maudemarie Clarke and Alan J. Swensen (Indianapolis: Hackett, 1998).
The Gay Science, trans. Walter Kaufmann (New York: Vintage, 1974).
Human, All Too Human, trans R. J. Hollingdale (Cambridge: Cambridge University Press, 1983).
Twilight of the Idols, trans. R. J. Hollingdale (Baltimore: Penguin, 1968).
Untimely Meditations, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1983).
Thus Spoke Zarathustra, trans. Walter Kaufmann (New York: Viking Press, 1966).
The Will to Power, ed. Walter Kaufmann, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage Books, 1968).

OTHER WORKS CITED

Gemes, Ken (2009). 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual' (this volume).
Janaway, Christopher (2006). 'Naturalism and Genealogy', in Keith Ansell Pearson (ed.), *A Companion to Nietzsche* (Oxford: Blackwell, 337–52).
Leiter, Brian (2001). 'The Paradox of Fatalism and Self-Creation', in John Richardson and Brian Leiter (eds.), *Nietzsche* (Oxford: Oxford University Press, 221–54).
Pascal, Blaise (2000). *Œuvres complètes*, ed. M. Le Guern (Paris: Gallimard, 2000).
Pippin, Robert (2006a). *Nietzsche, moraliste français: La conception nietzschéenne d'une psychologie philosophique* (Paris: Odile Jacob).
____ (2006b). 'Introduction', in Adrian Del Caro and Robert Pippin (eds.), *Thus Spoke Zarathustra* (Cambridge: Cambridge University Press).
Velleman, J. David (1992). 'What Happens When Someone Acts,' *Mind*, 101, 461.
Vivarelli, Vivetta (1994). 'Montaigne und der freie Geist,' *Nietzsche-Studien*, 23, 79–101.
Williams, Bernard (1994). 'Nietzsche's Minimalist Moral Psychology,' in R. Schacht (ed.), *Nietzsche, Genealogy, Morality: Essays on Nietzsche's On the Genealogy of Morals* (Berkeley: University of California Press), 237–47.

This page intentionally left blank

5

Nihilism and the Free Self

Simon May

I. INTRODUCTION

In this paper I claim that it is a necessary—but *not* a sufficient—condition for freedom in the sense in which Nietzsche valorizes it that one be a self in which a maximum number of drives of maximal power is organized into an evolving hierarchy—a hierarchy that enables one to 'promise' oneself with respect to one's self-legislated values and that one can affirm with satisfaction.

There is, I suggest, another key condition for attaining Nietzschean freedom—oddly neglected by at least Anglo-Saxon scholarship—and one to which reflection and self-consciousness, especially of an historical nature, are central: such a self is also one whose life is a coming-to-grips with the challenge of overcoming nihilism in its full-blooded sense—that is, of overcoming Western history as the history of the 'will to nothingness'.

Nihilism in its full-blooded sense is more fundamental than a nihilism of indifference towards or rejection of our hitherto highest values, or the conviction, whether despairing or defiant, that there are no stable values that fill life with meaning or at least none that can be attained. Full-blooded nihilism is to will—often passionately—what is nothing. This is the (far from necessarily indifferent or despairing) will that, as Nietzsche characterizes it, Platonism–late Judaism–Christianity and, in general, 'slave morality' manifest. They will 'nothing' because they are driven by an all-consuming will to escape a world of suffering, a will that, because it repudiates what is constitutive of living—the loss or elusiveness of what we most desire, such as loved ones, health, achievements, predictability, joy, and ultimately life itself—wills what is not human life, not the world of transience, chance, fate, and time in which we are actually situated. In refusing to affirm that life is structured by the possibility of loss they imagine an ideal order out of which this possibility has been conceptually airbrushed, an order that is clearly not the one into which humans are born.

By contrast to the nihilist, the Nietzschean 'free spirit' can master, and live joyously in, the world as it is, regardless of its suffering. He does not need to

strive for a world cleansed of suffering, let alone to set this up as life's supreme ideal. Indeed he regards this as a preposterous idea that diminishes human life, the flourishing of which demands hardship and suffering. But because he too is the inheritor of Judeo–Christian–Platonic morality, and to that extent its product, he cannot be free (of the need for ideals that promise to banish or redeem suffering) without mastering the history of that morality. This is not merely a cognitive challenge: the challenge of understanding what it is to lead a life structured by the ascetic ideal. It is also to take responsibility for that inheritance as one's most intimate possession. For only in that way might one eventually liberate oneself from it and so win genuine freedom.

II. FREE SELFHOOD IN NIETZSCHE: A SKETCH

Let us return to the first condition, or complex of conditions, for freedom of the self in Nietzsche—that (a) a maximum number of drives, (b) each of them of maximal power, is (c) organized into an evolving hierarchy in which (d) one can take satisfaction. This condition proceeds from Nietzsche's conception of the self or 'soul' as a 'social structure of the drives and affects'.[1] Each of these affects intrinsically values:[2] there are goods it essentially seeks, as well as conditions it needs for successfully attaining those goods; and there are things it avoids or despises. These are its 'yes's and 'no's'—for itself and to the world. Cognitively, each of these major drives and the valuations it embodies also grasps the world from particular perspectives: the things it wants, the ways it values, condition how it sees the world. These concepts—of a drive, of a value, and of particular cognitive perspectives—cannot therefore be separated. A drive always values and cognizes; a value always expresses a drive (or drives); and cognition always occurs from the perspective of a value, a perspective which involves adopting or being receptive to a particular way of seeing the world.[3]

The more *effectively* the drives are ordered into a hierarchy—the more control the self has over itself and over the circumstances with which it is faced—the more it is autonomous. Hierarchy is a matter of commanding and obeying—of one's drives becoming organized in such a way that one is able to commit oneself to projects that matter to one. It is the sort of command that, say, the concert pianist seeks through years of practice—not just the command that enables him to achieve speed, precision, evenness, and rich tonal variety, but also

[1] BGE 12; cf. GM II 2

[2] For example: 'Moralities are . . . a *sign language of the affects*' (BGE 187).

[3] There is an analogous reciprocity between feelings and judgements or evaluations. For example: 'feelings are nothing final or original; behind feelings there stand judgments and evaluations which we inherit in the form of feelings (inclinations, aversions)' (D 35).

the command—at once over body, emotion, sensibility, and conception[4]—that enables his musical values, and the *Weltanschauung* that they in turn embody, to be expressed in his playing. There is clearly nothing inflexible about such hierarchy; on the contrary it is the indispensable prerequisite for freedom of expression: without years of striving for the most exhaustive control of oneself freedom in Nietzsche's (or any real) sense[5] is unattainable. Successful hierarchy is therefore not the result of something else called 'free will'; it is free will.

And, far from being static, it evolves. It evolves not just because it takes a great deal of hard work to 'promise oneself' in respect of, or to stand surety for, the aims and valuations that one's self embodies and legislates at any given time in one's life, but because those aims and valuations themselves evolve or are discovered in the process of living life, especially a life of experiment and risk, in which danger is not shunned. Such a person gets to know what he wants and needs in order to flourish—and is conscious of possessing the strength and discipline to do what it takes to fulfil those needs and wants. Willing is then free.

It is also *experienced* as free. This experience does not involve beliefs such as those that *ressentiment* fosters when, for example, it imagines that 'masterly' types are free to be weak—beliefs in a metaphysical account of agency that assumes a doer behind deeds[6] or our willing as *causi sui*.[7] The experience of freedom of the will that for Nietzsche counts as genuine freedom has at least three distinguishing marks: one delights in seeing oneself as a commander;[8] one delights in seeing one's command as effective[9]—in other words in seeing one's 'protracted and unbreakable' will in action, a delight that seems to involve intense awareness of one's own capacity to commit oneself to action over long periods and of how one stands in relation to others' capacities in this respect;[10] *and* (for the powerful priest, too, could experience such delights) one does not

[4] This is analogous to fighting, as Nietzsche says Goethe did, the 'mutual extraneousness of reason, senses, feeling, and will' and disciplining oneself to 'wholeness' (TI IX 49).

[5] e.g. 'The "free" man, the possessor of a protracted and unbreakable will'; 'this rare freedom, this power over oneself and fate' (GM II 2). Will is 'the affect of command'; freedom of the will is the degree of one's 'power of self-determination' (GS 347). Here Nietzsche is in a long tradition that sees self-discipline as a condition of freedom.

[6] A dualism that he famously repudiates (GM I 13).

[7] See, e.g. BGE 19 and 21.

[8] BGE 19; cf. GM II 2

[9] Such a person clearly 'identifies himself with the executor of the order' and doesn't just 'enjoy triumph over obstacles' but 'thinks within himself that it was really his will itself that overcame them'. This self-attribution of willing and of willing's successes, this way of seeing them as one's own, is Nietzsche says, what happens in every well-constructed and happy commonwealth (BGE 19, *passim*).

[10] 'The "free" man . . . possesses his measure of value: looking out upon others from himself, he honors and despises. . . . ' (GM II 2). Contemporary readers might find such passages hard to stomach, but this sense of comparing one's own capacity to promise and commit with that of others, and the extent to which one thereby gives and expects respect, is crucial to Nietzsche's account of agents' first person apprehension of their own freedom.

seek faith in a dogma that is given or collective (GS 347): one's will does not crave external guidance. Between passages such as these (BGE 19; GM II 2; and GS 347) Nietzsche provides the bare bones of a phenomenology of freedom, one that makes reference to a substantially different sort of self-understanding to an apprehension of freedom that depends on the idea of will as first cause. And this self-understanding is advanced by Nietzsche not simply as his 'take' on what the free individual might, if he happens to reflect on his experience of freedom, believe it to consist in; such self-understanding is, rather, a condition of being a free self. To put it crudely: a zombie that could unflinchingly fulfil all its promises with 'protracted and unbreakable' continuity would not, by that token, be free.[11]

Though the evolution of hierarchy demands ceaseless practice—the practice of commanding and obeying within oneself, the constant search for freedom through discipline—the attainment of its overall direction cannot be forced. It will take its own time, and depends on the emergence of what Nietzsche variously calls our innate 'organizing "idea" ' (EH II 9), 'single taste' (GS 290), and 'personal providence' (GS 277)—none of which have anything to do with a metaphysical 'doer' standing behind its 'deed'. Hierarchy—and its organizing idea—might be there all along, as part of one's inherited rank order of drives (A 57), or it might emerge as a victor from the on-going battles within the self (BGE 19). These battles are largely unconscious: for consciousness and concepts, though they have a vital role to play in understanding the life-denying functions of traditional 'morality' and so in a fostering of a 'revaluation of values', are also seen by Nietzsche as merely the spokesmen, or consequences, of the truces, victories, and defeats in the wars being fought subconsciously.

In any event, there is nothing more dangerous than to seek one's organizing idea too soon. As I have argued elsewhere,[12] premature attempts to discover and understand it would arrest the development of two crucial preparatory phases: first, the innocent, unimpeded experience that comes from forgetting, misunderstanding, and focusing oneself; and, second, the development of '*single* qualities and fitnesses' free from the straightjacket of overall goals or meanings (EH II 9). The path to 'becoming what one is'[13] is usually not straight, and involves blunders and wrong turns, 'wasted' energy, protective misunderstandings, and, occasionally, self-narrowing. Hence, we must keep the 'whole surface of consciousness . . . clear of all great imperatives'; be alert to the 'many dangers

[11] See Sebastian Gardner's paper in this volume for a much deeper analysis of the first person perspective in Nietzsche's philosophical project.

[12] May (1999: 109–10 and 191–3).

[13] One of Nietzsche's major aims, expressed in the subtitle to EH as well as in formulations such as these: '*What does your conscience say*?—"You shall become the person you are." ' (GS 270); or: 'We . . . *want to become those we are*—human beings who are new, unique, incomparable, who give themselves laws, who create themselves.' (GS 335).

that the instinct comes too soon to ''understand itself'' '; and '[b]eware even of every great word . . . [and] pose' (EH II 9). It is for such reasons, I think, that Nietzsche approvingly cites Goethe's opinion that 'Truly high respect one can have only for those who do not *seek* themselves.' (BGE 266); or: 'To become what one is, one must not have the faintest notion *what* one is' (EH II 9).[14]

A vitally important expression of one's single taste, or organizing idea, is giving ' ''style'' to one's character' or, if one cannot do that, then otherwise achieving what Nietzsche says is the one 'needful' thing: to '*attain* satisfaction with [one]self' (GS 290). To give style to oneself is to achieve a self-conception that is thoroughly satisfying, so that not merely our strengths but also our weaknesses appear to us as necessary to—as having a crucial role in—a complete self that we can wholeheartedly affirm.[15] Nietzsche also expresses this point by calling on us to be 'poets of our lives', and emphasizes that this is first of all about giving artistry to 'the smallest, most everyday matters' (GS 299). That the whole self is pleasing to the individual helps to release her power and encourage her towards the actions for which she is predisposed and of which she is capable. '[D]elicate reverence' for oneself belongs to a 'noble soul' (BGE 265; cf. BGE 263, 287); whereas someone who is dissatisfied with herself, even if she is otherwise a well-functioning individual, will be 'continually ready for revenge' (GS 290). And revenge, as we know from Nietzsche's presentation of *ressentiment* in the first essay of GM, is a great obstacle to the attainment of freedom. In sum: this self-reverence or self-satisfaction is anything but smugly complacent; it is, on the contrary, the quiet confidence needed to risk, and promise oneself to, difficult tasks—the bright, daring mood that is the very opposite of those most dangerous spirits: the spirits of revenge and of gravity.

Finally there are two other key elements to free selfhood in Nietzsche's thought: the *number* of drives that the self incorporates and their respective *power*. The second of these—power—is clear: the stronger each drive the potentially stronger, and so more sovereign over circumstances, is the whole—if the drives can be organized to work with, rather than against, each other.[16] The first point is more interesting: the more drives one has, and the more diverse they are, the more valuations and perspectives one can bring to bear on one's life, and

[14] Nietzsche expresses a similar idea in this passage: ' ''My thoughts'', said the wanderer to his shadow, ''should show me where I stand; but they should not betray to me where I am going. I love my ignorance of the future and do not wish to perish of impatience and of tasting promised things ahead of time'' ' (GS 287).

[15] 'It is practiced by those who survey all the strengths and weaknesses of their nature and then fit them into an artistic plan until every one of them appears as art and reason. . . . Here a large mass of second nature has been added; there a piece of original nature has been removed. . . . In the end . . . it becomes evident how the constraint of a single taste governed and formed everything large and small' (GS 290).

[16] When Nietzsche asks 'What is good?', he answers: 'Everything that heightens the feeling of power in man, the will to power, power itself' (A 2). And in the *Nachlass* he adds: 'The ascent on this scale [of force] represents every rise in value . . . ' (WP 710).

so the richer are the ways in which one can see and create. A philosopher of the highest sort, Nietzsche says,

> would be compelled to find the greatness of man, the concept of 'greatness', precisely in his range and multiplicity, in his wholeness of manifoldness. He would even determine value and rank in accordance with how much and how many things one could bear and take upon himself, how far one could extend his responsibility. (BGE 212)[17]

But this richness comes at a price—which can be high: for the greater the number and diversity of drives, the harder it is to organize them into a well functioning hierarchy, and the more numerous—and therefore, taken together, improbable—are the conditions the individual needs in order to thrive. For each drive, as we have said, has its own valuations, and so each drive has its own 'physiological demands for the preservation of a certain type of life' (BGE 3), which is perhaps Nietzsche's most concise definition of a value.[18] This might be why he so often stresses that the most richly endowed people—his 'higher men'—can also be the most vulnerable; for they are dependent on the most numerous conditions for flourishing. Their 'complicated conditions of life can be calculated only with great subtlety and difficulty', so that 'the higher the type of man that a man represents, the greater the improbability that he will turn out *well*' (BGE 62; cf. BGE 274, 276).

Now let us say that this, roughly, is Nietzsche's picture of the maximally free, autonomous self—the self he most values: such a self has the maximum number and diversity of drives, each of them maximally powerful and with its sustained yes's and no's, organized into a clear and aesthetically pleasing hierarchy by an organizing idea or single taste, which has the commanding strength to commit the individual to her chosen courses—i.e. to 'promise herself'. Such a self is 'free'; it can commit itself unflinchingly.

III. WAGNER AS CASE STUDY OF BECOMING A FREE INDIVIDUAL

One of the most remarkable illustrations of such a free self and of its tortuous development that Nietzsche offers us is to be found in the third of his *Untimely*

[17] The difference, then, between the 'free spirit' and the unfree is that the free seeks power over, and enjoys being in, the world as it really is—it has the hardness and the courage to look clearly at the world and let reality in all its manifoldness and, at times, unpleasantness 'hit it'. Only thus can it be creative—by, for example, transforming this raw material of unpleasant reality into the simplifying and life-enhancing illusions of art and systematic thought. The sort of 'falsification' that arises from strength, that wants to see the world as it is, is based on a deep assimilation of truth. Great untruths are, as it were, built upon great truths.

[18] In the *Nachlass*, Nietzsche provides an interesting, if cumbersome, expansion of this definition: 'The standpoint of "value" is the standpoint of conditions of preservation and enhancement for complex forms of relative life-duration within the flux of becoming.' (WP 715).

Meditations: 'Wagner in Bayreuth'. In fact, Nietzsche's description of Wagner's development from a seeming dilettante to someone who could seize his whole age and its rich but agonizing history with one commanding vision is doubly interesting because, as he says in *Ecce Homo,* 'at bottom' the conditions to which this essay bears witness 'speak only of *me*'; 'this essay is a vision of *my* future'.[19] The following passage hints at how an 'organizing idea' emerges from the fiendish complexity of a great person's nature, nurture, and life circumstances:

> The dramatic element in Wagner's *development* is quite unmistakable from the moment when his ruling passion became aware of itself and took his whole nature in its charge: from that time on there was an end to fumbling, straying, to the proliferation of secondary shoots, and within the most convoluted courses and often daring trajectories assumed by his artistic plans there rules a single law, a will, by which they can be explained. . . . But there was also a pre-dramatic era in Wagner's life, that of his childhood and youth, and one cannot pass this era in review without encountering riddles. He *himself* does not yet seem to be present at all and that which, with hindsight, one might perhaps interpret as a sign of his presence appears at first as the simultaneous existence of qualities which must excite misgivings rather than hopeful anticipation: a spirit of restlessness, of irritability, a nervous hastiness in seizing hold upon a hundred different things . . . (UM IV 2).

As Wagner at length achieves 'spiritual and moral maturity', two powerful drives crystallize out from all the seeming chaos: the one a 'vehement will', 'dark, intractable, tyrannical'—a will that would easily destroy someone with a 'narrow spirit'; the other drive 'creative, innocent, illuminated'. By good fortune, the two drives are 'loyal' to each other, despite their individual struggles for expression, and their frequent 'hostility when they crossed one another'. 'In the surrender of one to the other', Nietzsche observes, 'there lay the great necessity which had to be fulfilled if he was to be whole and wholly himself.' Yet, 'at the same time', 'it [this surrender] was the only act that did not lie in his own power, which he could only watch and endure' (UM IV 3 *passim*). This discovery of one's 'necessity' through surrender dovetails with Nietzsche's insistence, to which I referred above, on the need for a certain passivity (not the same as 'reactivity') and patience in letting our ruling idea emerge—the need to keep the 'whole surface of consciousness . . . clear of all great imperatives' and '[b]eware even of every great word . . . [and] pose' (EH II 9).

Through all this, Wagner applied what Nietzsche calls his extraordinary '*talent for learning*' (UM IV 3) in order to 'appropriate to himself all that is highest in culture'. Such a voracious capacity for learning is something that Nietzsche attributes to other geniuses, notably Goethe,[20] and he clearly sees it as vital to

[19] EH *The Untimely Ones*, 3; my italics.

[20] See, for example, TI IX 49–51. One measure by which Nietzsche evaluates people is how much truth they can face—in exactly this manner of seeing and learning as much about brute reality as one can face.

the development of the sovereign individual.[21] But such a mass of learning and of clear-seeing can crush, unless it is structured by thought (and/or counterbalanced by art).[22] And Wagner had, in turn, to control this mass of experience and knowledge by an ever firmer 'arch of thought', so that it 'did not stifle his will to action' or 'entice him aside'.

Despite its overwrought prose, *Wagner in Bayreuth* is a fascinating case study of the development of Nietzsche's highest type of person; of their *Auseinandersetzung,* through constant trial and error, with their own many-sidedness; of how, step by step, they become sovereign over the competing elements within them as well as over external circumstances; of how such a life becomes ever more perilously complicated and must be brought under ever tighter organization, yet without stifling or side-tracking its further development; of how an individual's ruling idea slowly and unpredictably appears; and so of how their personal fate becomes manifest. Nietzsche might have changed his mind about Wagner later in his life—though to the end he saw his 'intimate relationship' with the composer as 'by far the most profound and cordial recreation' of his life (EH II 5); yet, as a sustained illustration of how a great man struggles to find his voice, this essay is without peer in the Nietzschean corpus.

IV. THE SELF AS HISTORICAL

There is, however, one major problem with the picture of the free Nietzschean self that I have sketched so far, and also with its illustration in the early essay on Richard Wagner: it does not yet situate that self—indeed, situate it self-reflectively—in time, in history, in the thoroughgoing sense which the later Nietzsche believes to be essential. Abandoning a metaphysical worldview does not just mean abandoning absolute standards or sources of value, such as God, Reason, or The Good; it also means seeing life as inescapably in time, as structured by time, with no timeless realm to give it meaning. There is a sense in which time is the ultimate object of *ressentiment*, an even more fundamental

[21] For example: 'We . . . *want to become those we are*—human beings who are new, unique, incomparable, who give themselves laws, who create themselves. To that end we must become the best learners and discoverers of everything that is lawful and necessary in the world . . .' (GS 335). The human beings described in the first part of the sentence are clearly the sort of sovereign, responsible, free individuals alluded to in GM II 2.

[22] Art, in so far as it is beautiful and enables us to falsify reality with a good conscience, actually helps us to *sustain* our truthfulness, where honesty alone—especially about the pervasiveness of error—would 'lead to nausea and suicide' (GS 107). This, Nietzsche says, is the ground of our *'ultimate gratitude to art'* (ibid.). His claim, at least in his later works, if not yet in *The Birth of Tragedy*, is not that we need art in order to avoid truth altogether, or as an alternative form of life to one that values truth, or to aestheticize existence by seeing it 'as' a work of art, but, exactly to the contrary, that we need art in order really to be able to face the truth, in order to seduce to life those who have been shaken by genuine insight into its horrors—'lest we *perish of the truth*' (WP 822).

object, perhaps, than the 'masterly' strength or flourishing of other people. As Zarathustra says in one of his more lucid utterances: 'This, indeed this alone, is what *revenge* is: the will's ill will against time and its "it was" ' (Z II 20 'On Redemption'). Even those with a strong will, it seems, those who might have liberated themselves from slavish *ressentiment* against the 'masterly' or the more fortunate, are not by that token alone free of the desire to take revenge against time:

> Will—that is the name of the liberator and joy-bringer; thus I taught you, my friends. But now learn this too: the will itself is still a prisoner. Willing liberates; but what is it that puts even the liberator himself in fetters? 'It was'—that is the name of the will's gnashing of teeth and most secret melancholy. Powerless against what has been done . . . [t]he will cannot will backwards; and that he cannot break time and time's covetousness, that is the will's loneliest melancholy (ibid.).

A free self—a self that is not merely an effective hierarchy of drives but is also not structured by metaphysical notions, by the ascetic ideal—must, therefore, be a self that affirms itself as a product of its history, a history that is a brute given and that can never be undone or be otherwise. As Zarathustra again puts it: Redemption is 'to re-create all "it was" into a "thus I willed it". . . . All "it was" is a fragment, a riddle, a dreadful accident—until the creative will says to it, "But thus I willed it . . . thus shall I will it" ' (ibid.).

In other words, if there is a place after the 'death of God' for the old religious word 'redemption', it is as release from *resentment against* time-boundedness. Where redemption used to be thought of as release from time-boundedness itself, it is now seen as willing temporality. Where redemption once involved justifying the evil of this temporal world in terms of a higher and timeless good to emerge from and vindicate suffering, redemption is now to be thought of as affirming transience and the suffering it might bring—to 'achieve a reconciliation with time, and something higher than any reconciliation' (ibid.). Since, Nietzsche holds, each of us is a being determined by our specific past, to affirm temporality entails affirming, *without justifying*, the 'piece of fatefulness' that we each are, its necessity and its beauty. And such affirmation of our fatefulness, such willing of the 'necessity'[23] of our own individual nature and deeming it beautiful, is, of course, precisely what Nietzsche takes '*amor fati*' to be (GS 276).

A self that expresses 'freedom of the will'—that is, mastery over itself and thus over its circumstances—is able to will its own necessity in this way, to recreate

[23] BGE 213. See also TI V 'Morality as Anti-Nature'; TI VI 'The Four Great Errors', especially section 8, where Nietzsche talks about the 'fatality of [one's] essence', that '[o]ne is necessary, one is a piece of fatefulness'; GM P 2 where Nietzsche says that ' . . . our values, our yeas and nays . . . grow out of us with the necessity with which a tree bears fruit . . . evidence of *one* will, *one* health, *one* soil . . . ' (GM P 2); and BGE 231, where he says that our nature is that inner 'granite of spiritual *fatum*' which sets for us certain 'predetermined . . . questions' that guide our thoughts and actions, and which disposes us to 'predetermined . . . answer[s]'. This is also what Zarathustra calls 'my *own* necessity' (Z III 12 sect. 30).

the whole history out of which it arises as a 'thus I will it'. This is precisely the 'trusting fatalism' that Nietzsche attributes to Goethe (TI IX 49; cf. BGE 39), his paragon of a free individual. The central faith of this fatalism is 'that all is redeemed and affirmed in the whole [person]' (ibid.)—a faith that denotes Nietzsche's later conception of 'Dionysian'.[24]

Yet there is a still more thoroughgoing affirmation of temporality—of a time-boundedness from which not even death finally releases us—than *amor fati,* namely to will the eternal recurrence of the same fate, to will infinite time for the specific 'piece of fatefulness' that one is (providing, I suggest, that this willing has 'strong' motives, which is not necessarily the case).[25] *Amor fati* is not as stringent a test as willing the Eternal Return because it is consistent with willing one lifetime and saying 'that's enough!'—an attitude that is still resigned, and which is still compatible with seeking an eventual exit from this world, from time, and from suffering. By contrast, affirming the Eternal Return is not consistent with welcoming any exit from time or suffering.

V. HISTORY AS NIHILISM

But what is the history—the 'it was'—in which any contemporary self is situated and with which it must reckon if it is to achieve freedom? The answer, as I suggested at the beginning of this essay, is this: our history—the history of Platonic–late Jewish–Christian cultures—*is* the history of nihilism. Nihilism isn't just one feature of Western history, let alone merely of an epoch in which men and women (1) ceased to believe in God and in traditional values and concepts associated with a dualistic moral world-order, and (2) could find no alternative, non-metaphysical ultimate meanings or purposes to believe in, or (3) experienced any such ultimate meanings or purposes that they did believe in as unachievable, or (4) positively believed that no such meanings or purposes exist or could exist. Nihilism is the character of Western history as a whole. The conclusion to GM presents the premisses for just this idea—which is explored

[24] Nietzsche's later conception of the Dionysian (e.g. TI IX 49; BGE 295) denotes the wholeness and multiplicity of an individual, his successful integration of a complex 'soul'—in contrast to his early conception in *The Birth of Tragedy*, where the god represents the formless, intoxicated, uncontrolled, innermost depths of nature, as distinct from the form-creating, controlled, art-making force of the Apollonian (e.g. BT 5–7 *passim*).

[25] There are, surely, potentially 'weak' motives to affirm Eternal Return—namely out of fear of death as a final stop to one's existence, or fear of irrecoverably losing one's past, or the desire for eternity. Affirming 'Eternal Return' or *amor fati* is a test of strength only if it involves affirming what we most dread. If what an individual most fears is not, say, the 'small man' (Zarathustra's obsessive dislike) or the recurrence of horrors, but, say, death and the absence of any eternity—just the sorts of fears that, according to Nietzsche, helped generate the life-denying metaphysics of transcendence—then affirming 'Eternal Return' will be easier for him than affirming a single non-recurring life. Indeed, the test of strength for such an individual would be to affirm a *finite* existence.

in much more detail in the notes gathered in Book 1 of WP. In section 28 of the third essay of GM Nietzsche says: (1) 'Apart from the ascetic ideal, man, the human *animal*, had no meaning so far . . . '. (2) 'We can no longer conceal from ourselves *what* is expressed by all that willing which has taken its direction from the ascetic ideal: this hatred of the human, and even more so of the animal, and more still of the material . . . this longing to get away from all appearance, change, becoming, death, wishing . . . all this means—let us dare to grasp it—*a will to nothingness*, an aversion to life, a rebellion against the most fundamental presuppositions of life . . . '. Therefore: (3) All meaning attributed to man, the human animal, so far, has taken its direction from the 'will to nothingness'.

We should not underestimate the central position attributed by Nietzsche to this presentation in GM of 'the will to nothingness'. Far from being advanced as part of 'a beginning that is *calculated to mislead*', with which he says in EH all the three essays of GM set out (EH on GM), it appears at the climax of the third essay, and so of the whole book. And when in EH he picks out the '*new* truth' to emerge from the third essay, he repeats that this new truth is the insight that the ascetic ideal is so tremendously powerful because it was 'the only ideal so far, because it had no rival'. And again he equates the ascetic ideal with willing 'nothingness' (EH on GM).

Now if the 'will to nothingness' is a fundamental feature of the history of which each of us scions of Platonism–Judeo–Christianity is a product, the task of willing our own necessity—and so of achieving individual freedom—must involve a profound reckoning with this pervasive will. Since revenge—the driving force of the ascetic ideal—is, according to Zarathustra anyway, ultimately directed against time itself, overcoming the will to nothingness, which takes its direction from the ascetic ideal, must ultimately involve overcoming revenge against time. And overcoming revenge against time, as I have just said, is most indubitably expressed by affirming the eternal return of one's own fatedness. In short, my conclusion, so far, is that affirming the eternal return out of the right motives would signal that one has overcome the will to nothingness—and in that sense nihilism itself.

For much of its history the will to nothingness has had manifestations that have been far from indifferent towards their values: on the contrary they have enthusiastically affirmed a conception of the world, structured by the ascetic ideal, that promises their adherents escape from suffering and all its putative causes; and they have denigrated a world-order that they experience as a cause of their suffering. They have therefore regarded life as anything but pointless and lacking ultimate values: life's meaning and highest ideal is to master and escape suffering—whether, as in religious conceptions, through a final transcendence or extinction of 'this world' or, as in secular conceptions, through a life devoted to the maximization of comfort and an ever more refined avoidance of hardship. They have, moreover, believed these values to be attainable, if only imperfectly,

or by the few who are blessed with the necessary abilities, or in an afterlife free of the constraints of this world, or in a society in which technology and legislation can, between them, progressively eliminate suffering and the causes of suffering.

By contrast, it is only when such a conception of life as ultimately deliverable from suffering seems unattainable, absurd and worthless, that we encounter nihilism in further senses of the term: indifference towards, or repudiation of, hitherto dominant values—on epistemological grounds (these values seem to be conceptually incoherent, or to refer to states of affairs or conditions for their fulfilment that do not or could not exist, or otherwise to make false promises) and/or on grounds of taste (they jar with prevailing norms of what seems 'decent' to believe). This indifference or repudiation might be coupled with lack of belief in an alternative value-system, either because an alternative has not been discovered yet (though one believes that such a system is discoverable in principle) or because one now believes that there are, and can be, no stable values that fill life with meaning or at least none that can be attained.

These further senses of nihilism appear only in certain historical phases (one highpoint was in nineteenth-century Europe). They are consequences of the desire for ultimate deliverance from suffering and depend on loss of faith in the possibility of attaining that end. And this desire is what nihilism is most fundamentally about—what the 'will to nothingness' most fundamentally expresses.[26]

Very schematically, one might say that the will to nothingness has expressed itself in three historical phases:[27]

Phase 1: Affirming the Unconditioned

In its first phase the will to nothingness affirms the unconditioned—be it called the 'God' of monotheism, the 'forms' of Plato, unconditionally valuable values (especially unconditional 'faith in truth'), or other such metaphysical categories. Affirming the unconditioned necessarily affirms another world because 'this world' is inescapably a realm of the conditioned in which all things are interconnected (e.g. WP 584). Since there is no (intelligible) realm of the unconditioned, to affirm it is quite literally to affirm what, for us, must be The Nothing.

[26] The most common understanding of what Nietzsche means by 'nihilism' is confined to these further senses of the term. Thus Bernard Reginster, who rightly places the diagnosis and overcoming of nihilism at the centre of Nietzsche's project, nonetheless defines the theme of nihilism as 'despair over the unrealizability of our highest values' (Reginster 2006: 51). But this despair is a late result of the more fundamental nihilism that wills nothingness, the nihilism that, as Nietzsche describes it, is at the heart of Platonism–Christianity and its secular successors and that neither takes its highest values to be unrealizable nor is in any despair over them. On the contrary.

[27] I revise here my proposal that the will to nothingness has four distinct phases, originally advanced in May (1999: 155–8).

Phase 2: Loss of Faith in God and/or a Moral World Order

Eventually, 'faith in truth' brings about the second phase of the will to nothingness: the discrediting, to greater (e.g. Hume) or lesser (e.g. Kant) degrees, of belief in God and in absolute external authority for moral values, for the social order, and for judgments of truth and beauty. This phase of nihilism is characterized by a strange mixture of demoralization and triumphalism. It is demoralizing because man's progressive naturalization undermines his former respect for himself and his allegedly supreme position in the 'order of creation' that religious cosmologies had inculcated; he becomes increasingly sceptical about the possibility of any intrinsic meaning or moral order in the universe—seeing even the concept of such meaning as redolent of the old metaphysical order—and so ends up 'slipping faster and faster . . . into nothingness', into a 'penetrating sense of his nothingness' (GM III 25). But man's release from external authority is also triumphalist because he now sees himself as the creator and arbiter of beauty, value, and even truth (GS 343, GS 344, GS 357, GM III 27)—as the legislator of everything that he formerly saw as given.

Despite the upheaval of the 'death of God', this phase is, ethically, remarkably conservative: man's core values remain stubbornly fixed—above all the traditional moral values of equality, pity, and truth. Indeed, these three central values of the morality that Nietzsche calls into question not only continue to be treated in people's practices and beliefs as if they were constitutive of the good; they are, if anything, espoused with *even greater* fervour than before the 'death of God'.

With *equality,* what happens is that a value that was once parasitic upon a Christian conception of souls as equal before God (or, equivalently, in the quality-less subject which, by definition, is equal to all other such subjects) is now transposed to the earthly domain, to relations between human beings, where it replaces the hierarchical social structures that start disintegrating in this same historical period of the 'death of God'. *'Liberté, Egalité, Fraternité'* expresses this well, as do feminism and all the ideals of modern democracy and egalitarianism, which obviously take as their premiss and goal the deep equality of human beings, regardless of their differences in ability or social roles or gender or wealth or ethnicity or culture.

Pity, too, according to Nietzsche, continues to be regarded as constitutive of the good long after we cease to believe in God, mainly due to our immense fear of suffering—a fear that drives us to seek all possible means not to experience it ourselves or to witness it in others. Indeed, once the consolations of Christian explanations and justifications of suffering are no longer available, pity becomes, if anything, more intensely valued than ever—a point to which I will return.

As to truth (and truthfulness), this continues to be treated as unconditionally valuable even—or precisely—by those whose piety for truth has caused them to repudiate a metaphysical conception of truth or of the authority for values

in general (GS 344). That faith in truth discredits theology or metaphysics in this sense, but not key values originally justified by it, cannot, I take Nietzsche to suggest, be an oversight, but must flow from the same fundamental will that takes 'its direction from the ascetic ideal', namely 'a will to nothingness' (GM III 28). This will is sustained despite the massive scepticism directed by modernity against all its metaphysical expressions and presuppositions—sustained simply because no alternative meaning for life to one structured by the ascetic ideal has yet been sought.

Phase 3: Devaluation of 'Morality's' Highest Values

In the third phase we see the undermining of these and other highest values that had so far remained sacrosanct despite the loss of belief in metaphysical justifications for them. This can go all the way to regarding all prevailing values as worthless and to believing in nothing (or imagining that one believes in nothing). 'One interpretation has collapsed', Nietzsche says in his *Nachlass*; 'but because it was considered *the* interpretation it now seems as if there were no meaning at all in existence, as if everything were in vain' (WP 55).

Of all the phases of the will to nothingness this one engenders the most despair and confusion. There is here a whole continuum of suspicion that extends from radical mistrust of specific past values (and contempt for man for believing in them) to the conviction that no alternative ultimate values exist or can exist, even in principle, let alone in practice, and from there to doubt that anything can have stable value—that any value can resist sceptical inquiry.

Above all, the value that has powered the progress of nihilism until now—the value of truth and truthfulness—itself comes into question. This value has cast fatal doubt on the old metaphysically-structured order—on the existence of God and, more widely, on the attainability of a good that supposedly transcends and redeems an imperfect world—and has, in the process, made man seem increasingly insignificant by disposing of so many former buttresses to his self-respect that were rooted in a metaphysical world-order. Though the will to truth makes man's self-contempt worse by exposing the astoundingly complex ramifications of our traditional metaphysical delusions, the full recognition and repudiation of which is the real meaning of the 'death of God', it nonetheless drives those 'heroic spirits who constitute the honor of our age' (GM III 24) and who cling to 'faith in truth' and to 'intellectual cleanliness at any price' (GM III 27; GS 357) to ask the key question—what is the meaning of the will to truth *itself*?—and to begin to draw the fatal inferences against this, their own highest, value.

Specifically, unconditional faith in truth calls into question its own value both by recognizing that truth is not always valuable (let alone that its value always trumps all other values) and by forcing us to see that the idea of the unconditioned is itself parasitic upon the unintelligible idea of timeless standards

of value, and so is yet another expression of the will to nothingness. Once this happens—once the 'will to truth gains self-consciousness'—'morality will gradually perish . . . : the great spectacle in a hundred acts reserved for the next two centuries in Europe' (GM III 27).

This drama of the perishing of morality is the end-point of what Nietzsche describes in his notebooks as the essence of nihilism: 'the highest values devalue themselves' (WP 2) under pressure of the will to truth. For these highest values, especially unconditional faith in truth, affirm an order of things which they also end up discrediting. The value of truth and truthfulness is, in this sense, the last of the highest values to be devalued.

IV. OVERCOMING NIHILISM AND ACHIEVING THE FREE SELF

There is, however, one small piece of unfinished business that remains, even after a hundred acts worth of drama: man's contempt for himself. Until man has overcome this self-contempt he has not overcome the legacy of the will to nothingness. As a result he cannot be truly free.

The key to overcoming self-contempt is 'to recreate all "it was" into a "thus I willed it" ', so achieving something like Goethe's 'trusting fatalism'. At first, this attitude seems at odds with Nietzsche's project: after all to revalue all values, to oppose the entire history of morality, seems anything but willing the past.

But mouthing formulae, such as Zarathustra's, isn't of course the point. What is crucial is *how* we will our fatedness. We are still stuck in nihilism if we will it with a shrug or between gritted teeth, out of a reluctant recognition that we have no alternative but to inherit the history that we inherit. We are likewise stuck in nihilism if we perform a calculus of welfare or look for a theodicy that tells us that our fate is, on balance, for the good, even if it is not for the (theoretical) best. And we are embroiled in nihilism if we persist in judging the value of wholes over against which we putatively stand—wholes such as 'the world' or 'existence' (GS 346) or 'time' or 'man, the human *animal*' (GM III 28). For, phenomenologically, each of these wholes designates a realm over against which we claim a position—in terms of our experience of them—as extrinsic judgers, a position from which we make overall assessments of their nature and value. To do this, to succumb to this posture, is to avoid our situatedness in them, to avoid living them. It is yet another way of imagining that we can distance ourselves from—and somehow grasp, as either knowers or evaluators—those dimensions of life whose brute existence makes us feel most helpless and therefore gives rise to the most recalcitrant suffering.

To overcome nihilism is to cease judging the value of such wholes at all, and to cease structuring the world and value-systems by reference to them. It is to affirm fate without qualification, and without even being tempted by thoughts

of its overall value. Perhaps only Nietzsche's putative 'free spirit'[28] no longer buttresses his dignity by ascribing to himself a capacity for evaluating a whole that he finds wanting, no longer seeks the intoxication that comes from such panoptic judgements—including, perhaps, those judgements about the whole history of morality that Nietzsche's philosophical project seeks to make.

But again we need to proceed carefully. To affirm fate, without qualification, is not simply an affective-cognitive state, however sincere and committed. It demands more action than that.

Let us recall that, at least for an inheritor of Platonic–Judeo–Christian morality, affirming fate specifically involves overcoming *revenge* against time. To affirm our fate—and our time-boundedness—means willing necessity not only backwards *but also* forwards. One wills backwards—one wills one's heritage—by turning all 'it was' into a 'thus I willed it' (Z, II 10). And one wills forward by 'promising oneself' (GM II 2) in respect of the values with which this historical determination endows one—by being strong enough, in practice, to maintain one's word 'in the face of accidents, even "in the face of fate" ' (ibid.).

The capacity to promise oneself in the fullest sense—in other words in respect of one's own sovereignly legislated values, those values that our fatedness has, as it were, built into us and whose necessity we will—is the mark of the truly free and sovereign individual (GM II 2), of 'the man of the future who can redeem us from the ascetic ideal, the will to nothingness, nihilism' (GM II 24). If we are not to live at the expense of the future (as Christianity did), we must develop the capacity to 'stand security' for it—'which is what one who promises does' (GM II 1). Nietzsche's words are '*für sich als Zukunft gut sagen zu koennen*'—to be able to promise oneself as a future, or 'as future' (ibid.). In other words, to be able to promise, to be responsible, *is* to live fully in time—to *be* one's future.

Morality—above all the 'morality of pity'—stands in the way of such a future and of real self-responsibility. Why does Nietzsche single out pity over other values of morality as the danger of dangers to genuine individuality? He clearly does so in, for example, GM P 5: 'it was precisely here' [in the 'instincts of pity, self-abnegation' and the 'un-egoistic'] he says, 'that I saw the *great* danger to mankind, its sublimest enticement and seduction—but to what? to nothingness? . . . .I understood the ever spreading morality of pity . . . as the most sinister symptom of a European culture that had itself become sinister, perhaps as its by-pass . . . to a Buddhism for Europeans? to—*nihilism*?'. Elsewhere he says that pity 'is the only religion preached now' (BGE 222). The role of pity is

[28] '. . . one could conceive of such a pleasure and power of self-determination, such a *freedom* of the will that the spirit would take leave of all faith and every wish for certainty, being practiced in maintaining himself on insubstantial ropes and possibilities and dancing even near abysses. Such a spirit would be the *free spirit* par excellence' (GS 347).

therefore not 'an isolated question mark'; to ponder it *is* to raise, in its acutest possible form, the question of the value of morality as a whole (GM P 6).

The reason for Nietzsche's emphasis on the morality of pity is, I suggest, because it—perhaps more than any other traditional value—attacks the precondition of freedom: to will and take responsibility for who we are, for a fate that is inevitably full of suffering. The morality of pity is, as the will to truth gradually reveals, deeply untruthful about the pitied—about *who they are* and *what they need*. Because (moral) pity—the pity that is driven by a desire to eliminate, at all costs, suffering and the sight of suffering—strips away from others' suffering whatever is distinctively personal, it in effect *lies* about who they are (GS 338). And by helping others indiscriminately to misinterpret or anesthetize their suffering, pity enables them to avoid the 'only . . . discipline [that] has created all enhancements of man so far' (BGE 225; cf. BGE 270; TI IX 38), the discipline required to breed 'an animal with the right to make promises'.

In other words, if our morality bids us minimize suffering—and even bids us engage in the fantasy that suffering can be 'abolished' (BGE 225; cf. 202)—then that morality will not just obstruct, but will, in effect, *prohibit* us from achieving responsibility for our own nature and values and experiences, and so from attaining free selfhood.

And yet nothing is more tempting than to pity others, to lose oneself in eliminating their suffering, so as not to focus on mastering one's own suffering and one's own path. Nietzsche himself confesses to finding that temptation overwhelming (GS 338). 'All such arousing of pity . . . is secretly seductive', he says; 'for our "own way" is too hard and demanding and too remote from the love and gratitude of others, and we do not really mind escaping from it . . . —to flee into the conscience of others' (GS 338 *passim*).

This is why 'a man who is by nature a *master*' (BGE 293) *has* got a right to pity. Pity will not undermine either him or the pitied; it will be a call to self-responsibility, not a draining of it; so that 'when such a man has pity . . . *this* pity has value' (BGE 293).

VII. CONCLUSION

My discussion of nihilism as the will to nothingness suggests that in order to understand Nietzsche's conception of the free self we cannot picture it simply as a hierarchy of diverse and sometimes competing drives, which is forged by an 'organizing idea', guided by a 'personal providence', gilded by giving it 'style', and expressive of one's own personal values. This would be altogether too static a picture of the self. Though such a self would be quite an accomplishment, it is no more than a necessary condition of freedom: a self structured by the ascetic ideal, or evincing a 'will to the end'—a 'priest', for example—might be highly organized in just this way, but will not be genuinely free.

To be free one must also achieve a fundamental reckoning with, and overcoming of, the will to nothingness as the will that runs though Platonic–late Jewish–Christian history and that makes it the history of nihilism. And to do this one must no longer rage against time: one must be able to will one's heritage and also to promise oneself in respect of one's values and the goals they express—in other words in respect of 'what one is'. Only thus has one prepared oneself to '*become* what one is'.

REFERENCES

WRITINGS BY NIETZSCHE

The English translations of Nietzsche's works to which I have referred are listed below.

The Antichrist (1888), trans. Walter Kaufmann, in *The Portable Nietzsche*, ed. Walter Kaufmann (New York, NY: Viking, 1954).

Beyond Good and Evil (1886), trans. Walter Kaufmann (New York, NY: Vintage, 1966).

Ecce Homo (1888), trans. Walter Kaufmann (New York, NY: Vintage, 1967).

The Gay Science (1882; Part 5: 1887), trans. Walter Kaufmann (New York, NY: Vintage, 1974).

On the Genealogy of Morals (1887), trans. Walter Kaufmann and R. J. Hollingdale (New York, NY: Vintage, 1967).

Thus Spoke Zarathustra (1883–1885), trans. Walter Kaufmann, in *The Portable Nietzsche*, ed. Walter Kaufmann (New York, NY: Viking, 1954).

Twilight of the Idols (1889), trans. Walter Kaufmann in *The Portable Nietzsche*, ed. Walter Kaufmann (New York, NY: Viking, 1954).

Untimely Meditations (1873–1876), trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1983).

The Will to Power, trans. Walter Kaufmann and R. J. Hollingdale (New York, NY: Vintage, 1968).

OTHER WORKS CITED

Gardner, Sebastian (2009). 'Nietzsche, The Self, and the Disunity of Philosophical Reason' (in this volume).

May, Simon (1999). *Nietzsche's Ethics and his War on 'Morality'* (Oxford: Oxford University Press).

Reginster, Bernard (2006). *The Affirmation of Life: Nietzsche on Overcoming Nihilism* (Cambridge, Mass.: Harvard University Press).

6

Nietzsche's Theory of the Will*

Brian Leiter

It is well-known that Nietzsche has provocative views about the nature of the will and free will.[1] It is less often appreciated that his views on these topics have considerable merit. Nietzsche not only anticipates and lends argumentative support to the new wave of non-libertarian incompatibilism defended by philosophers like Derk Pereboom (2001) and Galen Strawson (1994)—the view that free will is incompatible with "determinism" and that there is no credible account of free will as outside the causal order in the offing—but his theory of the will also wins some support from recent work on the will in empirical psychology (see Wegner 2002). As a philosophical naturalist, Nietzsche thought of his theoretical endeavors as proceeding in tandem with empirical inquiry (Leiter 2002: 6–11). As befits his self-designation as "the first psychologist," it turns out that Nietzsche anticipated results that psychologists only arrived at a century later.

In section 124 of *Daybreak*,[2] Nietzsche sets the primary issues that shall occupy us here in trying to understand his theory of the will. Nietzsche writes:

> We laugh at him who steps out of his room at the moment when the sun steps out of its room, and then says: "*I will* that the sun shall rise"; and at him who cannot stop a wheel, and says: "*I will* that it shall roll"; and at him who is thrown down in wrestling, and says: "here I lie, but *I will* lie here!" But, all laughter aside, are we ourselves ever acting any differently whenever we employ the expression "*I will*"?

* From Brian Leiter, "Nietzsche's Theory of the Will", *Philosophers' Imprint*, 7 (2007), 1–15.

[1] In discussing Nietzsche's "theory of the will," I am concerned with the notion of "will" familiar from general philosophy of action, both contemporary and historical, namely, the idea of a human faculty, whatever its precise character, that stands in some kind of necessary relationship with action. Such a faculty may itself be causally determined, or it may be autonomous of the antecedent causal order; its status may implicate questions of moral responsibility; and such a faculty may not exist at all. A theory of the will is one that sheds some light on these issues. My subject in this essay, I hasten to add, is *not* every use of the word "will" in Nietzsche's corpus, since those uses are quite various, and many have little to do with the familiar philosophical issues about the will just noted.

[2] I have generally started with the translations by Clark and Swensen, Hollingdale, and Kaufmann, and then made changes based on the Colli and Montinari edition of the *Sämtliche Werke*.

I take it to be uncontroversial that this last question is rhetorical, and that the intended answer to this question is: "no, it is no different at all." Should anyone be confused by the interrogative form Nietzsche employs (and, alas, some scholars have been so misled on other occasions[3]), other textual evidence will be adduced soon enough that makes it clear that when we act and say "I will," it is no different, and no more ridiculous, than when he "who steps out of his room at the moment when the sun steps out of its room . . . says '*I will* that the sun shall rise.' "[4]

If it is really true that this analogy holds, then it follows that the experience of willing which precedes an action does not track an actual causal relationship: the experience of willing is epiphenomenal (in some sense to be specified) with respect to the action. As Nietzsche notes in *The Gay Science*, "the feeling of *will* suffices for" a person "to assume cause and effect" (GS 127), but it will be the burden of Nietzsche's argument to show that this assumption is mistaken. Thus, an adequate account of Nietzsche's theory of the will and action will require us to get clear about three claims: first, the phenomenology of "willing" an action, the experience we have which leads us (causally) to conceive of ourselves as exercising our will (to say "I will"); second, Nietzsche's arguments that the experiences picked out by the phenomenology are not causally connected to the resulting action (or, alternatively, not causally connected in a way that would underwrite ascriptions of moral responsibility); and third, Nietzsche's account of the actual causal genesis of action. On the latter score, we shall turn to some recent work in empirical psychology that, in fact, vindicates precisely Nietzsche's skepticism that our "feeling" of will is a reliable guide to the causation of action.

THE PHENOMENOLOGY OF WILLING

Nietzsche recognizes that we often *feel* as if we are exercising free will, but he is unusual among philosophers in scrutinizing that experience, breaking it into its component parts. The resulting account is admittedly a *revisionary* one—it is not ultimately designed to *vindicate* the epistemic reliability of the feelings involved—but it is, Nietzsche claims, the correct account of those feelings. The

[3] See Owen and Ridley (2003: 70–1), who apparently confuse Nietzsche's rhetorical questions with genuine questions, perhaps because of a propensity to read Nietzsche too literally conjoined with inattention to the other kinds of textual evidence discussed in Leiter (1998: 2002), as well as in this essay. I have, in fact, changed my view about the nature of Nietzsche's epiphenomenalism, although not about his fatalism. I return to Owen and Ridley's criticisms, below, in the long note 11.

[4] Nietzsche himself makes the point clear just a few sections later in the same work, noting "that we really are not responsible for our dreams—but just as little for our waking life" (D 128).

key discussion comes in section 19 of *Beyond Good and Evil*, whose account of the phenomenology of willing bears quoting at length:

> [I]n every act of willing there is, to begin with, a plurality of feelings [*Gefühlen*], namely: the feeling of the state *away from which*, the feeling of the state *towards which*, and the feeling of this "away from" and "towards" themselves. But this is accompanied by a feeling of the muscles that comes into play through a sort of habit as soon as we "will," even without our putting "arms and legs" into motion. Just as feeling—and indeed many feelings—must be recognized as ingredients of the will, thought must be as well. In every act of will there is a commandeering thought,—and we really should not believe this thought can be divorced from the "willing," as if some will would then be left over! Third, the will is not just a complex of feeling and thinking; rather it is fundamentally an *affect [ein Affekt]*; and specifically the affect of the command. What is called "freedom of the will" is essentially the affect of superiority with respect to something that must obey: "I am free, 'it' must obey"—this consciousness lies in every will . . . .A person who *wills*—, commands something inside himself that obeys, or that he believes to obey.

Let us use an example to flesh out Nietzsche's account. Sitting at the computer, I wonder whether I should go downstairs to see what the children are doing. I "decide" to do so, and so begin to rise from my chair. I *feel* as if I have willed the movement: I feel the *moving away* from the desk and computer, the *moving towards* the door, and I *feel* the physical, or muscular, movement as well. Let us call all this complex of feelings, for ease of reference, "the bodily feelings."

These bodily feelings are not, however, sufficient for the experience of will: they are merely *qualitative*, merely the "raw" feeling of "away," "towards," of muscles contracting, limbs moving. We still need the "commandeering thought"—that's Nietzsche's label, which I will use in what follows—namely, the thought, "I will get up from the desk and go downstairs" or some suitable surrogate. But the bodily feelings and the commandeering thought are still not enough, according to Nietzsche, for the experience of willing. This is perhaps Nietzsche's key claim. For the experience of willing is, according to Nietzsche, essentially the *meta-feeling*—the "affect" Nietzsche calls it—of commanding. "Affect" is not being used here in its Freudian sense of psychic energy or "charge," but as something closer to feeling again. By the "affect of the command," Nietzsche means the *feeling* that the *thought* (i.e., the propositional content, such as "I will get up from the desk and go downstairs"[5]) brings about these other bodily feelings, i.e., of "away from," "towards," of, in a word, movement; and that this commanding is *who I am*. By *identifying* with the commandeering thought—by taking that to be "who I am" (on this occasion)—we *feel* superior, we experience this affect of superiority. So we have the experience of willing when the person identifies himself with a certain propositional content (the commandeering thought "I

[5] It might be more apt to put the propositional content in the imperative form, so that it does not sound merely like a prediction!

will get up from the desk and go downstairs") that one takes to be commanding the bodily feelings, i.e., the feeling that attach to the "towards," "away from" and the muscular sensations; and this identification produce the meta-feeling of superiority which is the feeling of willing. In short, one experiences willing when one *feels* as if the bodily qualia are obeying the thought, and that the commanding thought is who I am.

As Nietzsche recognizes, there is something paradoxical here, since, as he observes in the same section of *Beyond Good and Evil*:

> [W]e are, under the circumstances, both the one who commands *and* the one who obeys, and as the obedient one we are familiar with the feelings of compulsion, force, pressure, resistance, and motion that generally start right after the act of willing. On the other hand, however, we are in the habit of ignoring and deceiving ourselves about this duality by mean of the synthetic concept of the "I." (BGE 19)

In other words, talk of "I"—as in "*I* will go downstairs"—obscures an elemental fact: it's my body that is "obeying" my will, so "I" am also one who obeys, as well as one who commands. Of course, we don't experience it, or think of it, that way: we identify the "I" with the feeling of *commanding*, not the feeling of obeying. Thus, our experience of the "I," our identification of it with the commandeering thought, itself requires an explanation: *why* do we not identify ourselves with the *commanded* feelings and movements, why do we, instead, identify ourselves with the superiority of the *commandeering thought*?

Here is what Nietzsche offers by way of explanation in the same passage:

> [T]he one who wills believes with a reasonable degree of certainty that will and action are somehow one; he attributes the success, the performance of the willing to the will itself, and consequently enjoys an increase in the feeling of power that accompanies all success. "Freedom of the will"—that is the word for the multi-faceted state of pleasure of one who commands and, at the same time, identifies himself with the accomplished act of willing.[6]

We identify, then, with the *feelings of command* rather than those of obedience because identifying with the former increases pleasure (it is pleasant to feel the body "obeying")—so the explanation might seem straightforwardly hedonistic. But this would be too quick, since the real explanation, as often happens in Nietzsche, is cast in terms of feelings of power which, *in turn*, produce pleasurable

[6] What then of *unsuccessful* actions, which presumably one can also experience as being freely willed, although not "accomplished"? Suppose, for example, I decide to get up from the computer to see what the children are doing, but fail to do so (perhaps one of my kids has glued me to the chair!). Although the action is unsuccessful, all the requisite components for the experience of free will are present (e.g., I can feel the requisite bodily movements, even if they are not brought to completion). Trickier cases involve (as a referee pointed out) extreme fatigue or paralysis, though at least in the former case, it is not implausible that, as with the unsuccessful action, I can feel some distinctive muscular sensation, even if I do not move my limbs. Nietzsche's proposed phenomenology may not work for the paralyzed, but we would have to know more about what, if any, experience they really have of willing the motion of paralyzed limbs.

sensations. We need not resolve the issue here of the primacy of the desire for pleasure or desire for power as the fundamental explanatory mechanism;[7] for the phenomenology, all that matters is that it be true that there *is* a feeling of pleasure attendant upon the sensation of willing, *even if* that feeling derives from a feeling of power.

So now we have Nietzsche's account of the *phenomenology* of willing from *Beyond Good and Evil* (and I should add that I know of no more systematic account in the corpus that conflicts with it): we *feel* as though we are exercising free will when we *identify* with the "commandeering thought" which we feel is superior to, and being obeyed by, the myriad qualitative experiences involved in movement, that is, the bodily feelings; and we so identify because of the feelings of pleasure and power that arise from the "affect of superiority" that flows from that identification.[8]

WHY THE PHENOMENOLOGY DOES NOT TRACK A CAUSAL RELATIONSHIP

The crucial idea in Nietzsche's theory of the will is that the phenomenology of willing, no matter how vivid, does *not* in fact mirror or reflect or—as I shall say in what follows—*track* an actual causal relationship (or, more precisely, a causal relationship sufficient to underwrite ascriptions of moral responsibility, a crucial qualification to which we'll return). That is, the feeling of superiority that attaches to the "commandeering thought" with which we identify is not, in fact, identical with anything that actually stands in a causal relationship with the resultant action. As Nietzsche writes in the long section from *Beyond Good and Evil* that we have been scrutinizing:

> the one who wills [i.e., who has the experience of willing] believes, in good conscience, that willing *suffices* for action. Since it is almost always the case that there is will only where the effect of command, and therefore obedience, and therefore action, may be *expected*, the *appearance* translates into the feeling, as if there were a *necessity of effect*. In short, the one who wills believes with a reasonable degree of certainty that will and action are somehow one; he attributes the success, the performance of the willing to the will itself, and consequently enjoys an increase in the feeling of power that accompanies all success. . . . (BGE 19)

To put it in quasi-Humean terms, Nietzsche claims that because the complex of bodily feelings, commandeering thought, and meta-feeling are fairly constantly

[7] In *The Gay Science*, Nietzsche offers the apparently stronger claim on behalf of hedonism that for all willing "a representation of pleasure and displeasure is needed" (GS 127). Unfortunately, there is no elaboration of the thought, and so it is not clear that the *representation* might not itself be prompted by an experience of *power*: in any case, nothing in the discussion there rules it out.

[8] In what follows, I am going to take for granted the accuracy of Nietzsche's account of the phenomenology.

conjoined with succeeding bodily actions, we naturally infer (since it increases "the feeling of power") that the will has caused the subsequent actions. How, then, has the phenomenology misled us, according to Nietzsche?

Remember that the experience of will for Nietzsche has three components: the bodily feelings; the commandeering thought; and the meta-feeling of superiority, i.e., the feeling that the thought commands everything else. It is the meta-feeling (perhaps the pleasure attendant upon the meta-feeling) in turn that leads us to identify with the thought, rather than with the parts of our body that are commanded.

In debunking the phenomenology as a reliable guide to causation, Nietzsche's target is the *commandeering thought*, rather than the bodily feelings or the meta-feeling. And his argument is brilliantly simple. He starts from another bit of phenomenology, namely that, "a thought comes when 'it' wants, and not when 'I' want" (BGE 17). Nietzsche's target in this *particular* passage is the famous Cartesian doctrine, "I think, therefore I am." But, Nietzsche points out, from the fact that there is thinking, it does not follow that *I*, i.e., some subject or agent, is doing the thinking, and so it does not follow that I exist. As Nietzsche puts it:

> It is . . . a *falsification* of the facts to say that the subject "I" is the condition of the predicate "think." It thinks: but to say the "it" is just the famous old "I"—well that is just an assumption or opinion, to put it mildly, and by no means an "immediate certainty." In fact, there is already too much packed into the "it thinks": even the "it" contains an *interpretation* of the process, and does not belong to the process itself. (BGE 17)

Even if the explicit target in this particular passage is the Cartesian "I," the surrounding context makes clear the real target, namely, the will. The preceding section of *Beyond Good and Evil*, for example, treats the "I think" and the "I will" as both being common, but mistaken, examples of "immediate certainties;" while the following section returns explicitly to an attack on "free will," only to be followed by the long passage, section 19 of *Beyond Good and Evil*, on which we have been concentrating.

Now what does Nietzsche's phenomenological claim here mean? What does it mean to say a thought comes when "it" wants, not when "I" want? Because we are talking about thoughts that "come," I take it Nietzsche must mean thoughts that come to consciousness. And his point is that our "thoughts" appear in consciousness, without our having willed them: "*ein Gedanke kommt, wenn 'er' will, und nicht wenn 'ich' will*." We need to be careful here since, after all, Nietzsche is engaged in an attack on the existence of will, so he can't believe that there is any sense in which I could genuinely will a thought into existence. But he can't, of course, presuppose that conclusion here without begging the question. Rather, we must take the talk of willing here to refer to the *experience of willing*, which Nietzsche, as we have seen all along, concedes is real enough.

So Nietzsche's phenomenological point then comes to this: a "thought" that appears in consciousness is *not* preceded by the *phenomenology of willing* that Nietzsche has described, that is, there is no "commandeering thought" preceding the conscious thought to which the meta-feeling (the affect of superiority) attaches. (Even if there were such a commandeering thought in some instance, this would just create a regress, since not every commandeering thought will be preceded by the experience of willing.) Since we do not *experience* our thoughts as willed the way we experience some actions as willed, it follows that no thought comes when "I will it" because the experience to which the "I will" attaches is absent.[9]

Notice the rather clever structure of this argument, for its critique is entirely *internal* to the perspective of the agent who takes himself to possess a will. For what Nietzsche does is point out that the criterion of willing that agents themselves treat as reliable guides to a causal relationship—namely, the phenomenology described above—is, in fact, completely absent in the case of thoughts (or, at least, in the case of the thought that starts an inferential chain of thinking which involves the experience of willing). As an introspective matter, it seems to me Nietzsche is plainly correct about this point. But if we do not experience our thoughts as willed, then it follows that the actions that follow upon our experience of willing (which includes those thoughts) are not caused in a way sufficient to underwrite ascriptions of moral responsibility.

That conclusion came rather quickly, so let us pause for a moment, and try to set out more clearly how we got there. Nietzsche's conclusion—that our experience of willing does not, in fact, make us morally responsible for our actions—requires two premises. The first premise is this: one component of the experience of will—namely, the commandeering thought—is, itself, causally determined, at some point, by something other than the will, since thoughts that come into consciousness are causally determined by something other than the will.[10] (What the "something else" is we shall return to.) The second premise—the one so far unstated and unargued for in our discussion of Nietzsche's theory of the will—is that being self-caused (*causa sui* as Nietzsche calls it) is a necessary condition for responsibility (see BGE 21).[11] Since we have

[9] I agree with Katsafanas (2005: 11–12) that BGE 17 does not support the epiphenomenalitiy of consciousness *per se*, as I had wrongly claimed in Leiter (1998), but it does, as I argue here, support the epiphenomenal character of those experiences related to willing.

[10] I am going to speak of "caused" and "causally determined" interchangeably, without meaning to prejudge issues about probabilistic causation—issues that, in any case, were unknown to Nietzsche.

[11] Owen and Ridley object (2003: 73–4) that the rest of the passage (BGE 21) undermines the point noted in the text for two reasons. First, they write: "it must be a mistake to attribute to Nietzsche . . . the view that the will is 'unfree' in the sense of being causally determined, since he explicitly rejects that position" (73). But in what sense precisely does Nietzsche "reject" that claim? Owen and Ridley omit the relevant text which, as I discussed in Leiter (2002: 22–3), makes clear that Nietzsche's *apparent* skepticism has nothing to do with whether *the will is causally determined*,

shown that the "commandeering thought" that is part of the experience of will is not *causa sui*, it follows that the will it helps constitute is not *causa sui*, and

but with whether *anything* is causally determined: and on this score, Nietzsche is still, in BGE 21, in the grips of the NeoKantian view (acquired from his reading of Friedrich Lange) that "cause and effect" are merely features of the phenomenal world, not of "things in themselves." If Nietzsche abandons—as even Owen and Ridley acknowledge he does (2003: 74)—this NeoKantian view, then Nietzsche's argument against being *causa sui* is unaffected: if nothing in the "phenomenal" world can be self-caused, and the modifer "phenomenal" is doing no work, then it follows that nothing can be "self-caused" *simpliciter*. Owen and Ridley, in response to my comments on an earlier draft of their essay, purport to acknowledge this point, which brings us to their second objection, which warrants quoting (2003: 74), "[I]t does not follow . . . from Nietzsche's rejection of the *causa sui* that he instead either embraces or must embrace full-blown classical determinism, since, as Nietzsche makes clear, he regards the *causa sui* and classical determinism as symptoms of precisely the same "superlative metaphysical" nonsense . . . .[Even granting Nietzsche's rejection of Lange's Neo-Kantian skepticism concerning the reality of causation], if, as Leiter accepts, the argument against the *causa sui* as a piece of "superlative metaphysical nonsense" still holds once the Neo-Kantian view has been abandoned, then so too does the argument against classical determinism. That Nietzsche later came to accept the reality of causation, in other words, has no effect on the logical point that we draw from the passage, namely, that the opposite of a piece of nonsense is itself a piece of nonsense." This response is problematic at a number of levels. First, I certainly did not claim that Nietzsche embraces "full-blown classical determinism"—indeed, I am explicit about distinguishing classical determinism from Nietzsche's fatalism (1998: 224–5; 2002: 82–3)—though a separate misunderstanding (to which I'll return in a moment) may explain why Owen and Ridley think the latter collapses into the former. Second, Nietzsche does not speak of "superlative metaphysical nonsense," he says rather that the idea of the *causa sui* reflects a "longing for 'freedom of the will' in the superlative metaphysical sense," that is, in contemporary terms, some kind of libertarian free will. Even allowing that Nietzsche deems the idea of the *causa sui* nonsensical—he does say it "is the best self-contradiction that has even been conceived" even though he does not call it "superlative metaphysical nonsense"—it simply does not follow, logically or otherwise, that an opposed doctrine is also nonsensical, unless that opposed doctrine shares the premise that renders the first self-contradictory. But what makes the idea of the *causa sui* incredible is the idea that one can "pull[] yourself by the hair from the swamp of nothingness up into existence" (BGE 21), and *that* is precisely the idea that is repudiated by the view that the will is causally determined. Nietzsche's purported argument against this latter doctrine depends *entirely* on the NeoKantian skepticism about causation that even Owen and Ridley concede he later abandons. This, in turn, explains precisely why there was no point in quoting that portion of the passage: the argument against the "unfreedom of the will" is a bad argument, depending on a NeoKantian doctrine Nietzsche himself came to repudiate. By contrast, the argument against the *causa sui* is sound in its own right, independent of the NeoKantian skepticism, and an argument consistent with other claims Nietzsche makes in this and subsequent works. (For similar reasons, there is no reason to take particularly seriously Nietzsche's claim that "in real life it is only a matter of *strong* wills and *weak* wills," since that claim is motivated *entirely* in this passage by the NeoKantian skepticism regarding causation which Nietzsche abandons.)

As to why Owen and Ridley run together "classical determinism" and Nietzsche's fatalism (or "causal essentialism" as I call it), the answer is suggested by their claim that "causal essentialism [is] . . . the view that, at any given moment, *the* trajectory of a person's life . . . is determined by type-facts plus environment, a view which, given that type-facts plus environment just equals the totality of causally relevant facts, is indistinguishable from classical determinism" (2003: 74). Contrary to Owen and Ridley, causal essentialism (which they characterize correctly) is still distinguishable from classical determinism, since the latter requires that there be "laws of nature," and it is precisely because Nietzsche is a radical empiricist skeptic about *laws* that he eschews the language of classical determinism. Nietzsche's "official" view (strange as it may seem) is that ours is a world of *token* necessities, not lawful necessities, and that is true even if persons are of different *types*: the *types* interact with the environment in ways that are not law-governed.

thus any actions following upon that experience of willing could not support ascriptions of moral responsibility.

Now Nietzsche, it must be conceded, simply takes for granted that free will and moral responsibility are incompatible with causal determination of the will. Why he takes this incompatibilism for granted is not hard to see, though: it is deeply embedded in ordinary moral and religious thought. As Galen Strawson has noted, the incompatibilist idea of responsibility "has for a long time been central to the Western religious, moral and cultural traditions" (Strawson 1994: 8). Arguably, it is only certain philosophers who think the need to be a self-caused agent is superfluous, something that can be finessed via some adroit dialectical moves. (And even among philosophers, of course, discontent with compatibilism is legion: see, e.g., Nagel 1986: 113 for a striking example.)

The "argument from the phenomenology of thoughts" (as I shall call it)—namely, the argument that they come when they want, not when I want—is not, however, the only consideration influencing Nietzsche's conclusion that the will is causally determined. That is actually fortunate, since the argument from the phenomenology of thoughts doesn't rule out the possibility that our experience of willing misleads us as to our *real will*, and that this *real will* does, in fact, stand in the appropriate causal relationship with actions. Of course, this "real will," if it exists, had better be one that the agent can claim as his own if we are to then saddle the agent with responsibility for the actions it produces. But we will return to that issue below.

To see what the other argument for the epiphenomenality of the will might be, we need to remember—as I have argued elsewhere (Leiter 2002: 63–71)—that Nietzsche is very much influenced by the idea popular among German Materialists in the 1850s and after that human beings are fundamentally bodily organisms, creatures whose physiology explains most or all of their conscious life and behavior. Nietzsche adds to this Materialist doctrine the proto-Freudian idea that the unconscious psychic life of the person is also of paramount importance in the causal determination of conscious life and behavior.[12] Thus, Nietzsche, as I have argued, accepts what we may call a "Doctrine of Types" (Leiter 1998), according to which,

> Each person has a fixed psycho–physical constitution, which defines him as a particular *type* of person.

Call the relevant psycho–physical facts here "type-facts." Type-facts, for Nietzsche, are either *physiological* facts about the person, or facts about the person's unconscious drives or affects. The claim, then, is that each person has certain largely immutable physiological and psychic traits that constitute the "type" of

[12] Nietzsche's "official" view seems to be that physiology is primary, but he mostly concentrates on psychological claims, most obviously because he is no physiologist!

person he or she is. While this is not, of course, Nietzsche's precise terminology, the ideas are familiar enough from his writings.

A typical Nietzschean form of argument, for example, runs as follows: a person's theoretical beliefs are best explained in terms of his moral beliefs; and his moral beliefs are best explained in terms of natural facts about the type of person he is (i.e., in terms of type-facts). So Nietzsche says, "every great philosophy so far has been . . . the personal confession of its author and a kind of involuntary and unconscious memoir"; thus, to really grasp this philosophy, one must ask "at what morality does all this (does *he*) aim" (BGE 6)? But the "morality" that a philosopher embraces simply bears "decisive witness to *who he is*"—i.e., who he *essentially* is—that is, to the "innermost drives of his nature" (BGE 6). Indeed, this explanation of a person's moral beliefs in terms of psycho–physical facts about the person is a recurring theme in Nietzsche. "[M]oralities are . . . merely a sign language of the affects" (BGE 187), he says. "Answers to the questions about the *value* of existence . . . may always be considered first of all as the symptoms of certain bodies" (GS Preface 2). "Moral judgments," he says are, "symptoms and sign languages which betray the process of physiological prosperity or failure" (WP 258). "[O]ur moral judgments and evaluations . . . are only images and fantasies based on a physiological process unknown to us" (D 119), so that "it is always necessary to draw forth . . . the *physiological* phenomenon behind the moral predispositions and prejudices" (D 542). A "morality of sympathy," he claims is "just another expression of . . . physiological overexcitability" (TI IX 37). *Ressentiment*—and the morality that grows out of it—he attributes to an "actual physiological cause [*Ursache*]" (GM I 15). Nietzsche sums up the idea well in the preface to the *Genealogy*: "our thoughts, values, every 'yes,' 'no,' 'if' and 'but' grow from us with the same inevitability as fruits borne on the tree—all related and each with an affinity to each, and evidence of one will, one health, one earth, one sun" (GM P 2).

We can see Nietzsche's Doctrine of Types clearly at work in his theory of the will when we move from *Daybreak* and *Beyond Good and Evil* to another significant discussion of the will in the late Nietzschean corpus: "The Four Great Errors" section of *Twilight of the Idols*. Three of the errors Nietzsche discusses here pertain to causation: he calls them "the error of confusing cause and effect" (1–2), "the error of false causality" (3), and "the error of imaginary causes" (4–6). The fourth "great" error, by contrast, is "the error of free will" (7–8), though there is, in fact, no argument given in these concluding sections of the chapter for why free will is an error (instead Nietzsche offers one debunking explanation of why people might be motivated to believe in free will apart from its reality). The inference the reader is plainly supposed to draw is that the "error of free will" follows from the errors about causation discussed in the preceding sections. If we can reconstruct the argument that emerges from those sections,

then, we shall have identified Nietzsche's other main reason for viewing the will as epiphenomenal and thus for rejecting free will.[13]

The first error, that "of confusing cause and effect," can be summarized as follows: given two regularly correlated effects E1 and E2 and their mutual "deep cause," we confuse cause and effect when we construe E1 as the cause of E2, missing altogether the existence of the deep cause. Let us call this error "Cornarism" after the example Nietzsche uses:

> Everybody knows the book of the famous Cornaro in which he recommends his slender diet as a recipe for a long and happy life . . . I do not doubt that scarcely any book (except the Bible, as is meet) has done as much harm . . . .The reason: the mistaking of the effect for the cause. The worthy Italian thought his diet was the *cause* of his long life, whereas the precondition for a long life, the extraordinary slowness of his metabolism, the consumption of so little, was the cause of his slender diet. He was not free to eat little *or* much; his frugality was not a matter of "free will": he became sick when he ate more. (TI VI 1)

In other words, what explains Cornaro's slender diet *and* his long life is the same underlying fact about his metabolism. Cornaro's mistake was to prescribe his diet for all without regard for how individuals differed metabolically, metabolism being the relevant type-fact in this context.

Even if we grant Nietzsche all the facts as he presents them, this would not yet show that there is no free will, unless the error involved in Cornarism extended beyond cases such as diet and longevity. But that is exactly Nietzsche's contention, since in the very next section he saddles morality and religion quite generally with Cornarism. According to Nietzsche, the basic "formula on which every religion and morality is founded is: 'Do this and that, refrain from that and that—then you will be happy! Otherwise . . . .' " Cornaro recommended a slender diet for a long life; morality and religion prescribe and proscribe certain conduct for a happy life. But, says Nietzsche,

> [A] well-turned out human being . . . *must* perform certain actions and shrinks instinctively from other actions; he carries the order, which he represents physiologically, into his relations with other human beings and things.

So morality and religion are guilty of Cornarism: the conduct they prescribe and proscribe in order to *cause* a "happy life" are, in fact, *effects* of something else, namely the physiological order represented by a particular agent, one who (as Nietzsche says) "*must* perform certain actions," just as Cornaro *must* eat a slender diet (he is "not free to eat little *or* much"). That one performs certain actions *and* that one has a happy life are themselves both effects of the physiological order. If

[13] We shall return in the final section of the paper to the ambiguity regarding the strength of the epiphenomenalist doctrine as articulated so far.

we grant Nietzsche the Doctrine of Types, then there is indeed reason to think that Cornarism is a feature of morality too, since morality fails to recognize the crucial role of type-facts in determining what one does, even what morality one accepts.

That brings us to the next "error," that of "false causality," the mistake of thinking we know what causation is because of our introspective confidence in what we take to be the causal powers of our own mental life. Nietzsche explains:

> We believed ourselves to be causal in the act of willing: we thought that here at least we caught causality in the act. Nor did one doubt that all the antecedents of an act, its causes, were to be sought in consciousness and would be found there once sought—as "motives": else one would not have been free and responsible for it. Finally, who would have denied that a thought is caused? That the "I" causes the thought? (TI VI 3)

We already know, of course, from the phenomenology of thought argument who it is that denies that thoughts are caused by the "ego," by some internal agency, such that one would be "free and responsible" for them. And Nietzsche soon makes clear in this section of *Twilight of the Idols* that his view remains unchanged:

> The "inner world" is full of phantoms . . .: the will is one of them. The will no longer moves anything, hence does not explain anything either—it merely accompanies events; it can also be absent. The so-called *motive*: another error. Merely a surface phenomenon of consciousness—something alongside the deed that is more likely to cover up the antecedents of the deeds than to represent them. . . . .
>
> What follows from this? There are no mental [*geistigen*] causes at all. (TI VI 3)

In the last line, Nietzsche must mean only that there are no *conscious* mental causes (or, at least, no conscious mental causes of those actions with which morality is concerned).[14] Indeed, in other passages, he is explicit that the target of this critique is the picture of conscious motives as *adequate* to account for action.[15] As he writes in *Daybreak*, "we are accustomed to exclude all [the]

[14] Katsafanas (2005) is correct, I think, to object—*contra* Leiter (2002) as well as Gilles Deleuze—that Nietzsche cannot hold that consciousness *per se* is epiphenomenal, though I am ultimately skeptical about his own intriguing, often ingenious, presentation of Nietzsche's account of consciousness, since it floats free of the texts at crucial points. (It also makes the Freudian idea of unconscious wishes unintelligible [since their content must be conceptually articulated], though Freud was quite correct, in my view, to have found the germ of this idea in Nietzsche.) But there is plainly a need to individuate *those conscious* states whose causal efficacy is impugned by Nietzsche. Katsafanas's own interesting distinction between the epiphenomenality of the Ego (the "substantive faculty" of "Consciousness") versus causally efficacious conscious mental states (those mental states marked by their conceptually articulated content) (2005: 13) might be one way of marking the relevant distinction for Nietzschean purposes. Freud's account—in which those conscious mental states which conjoined an Idea and an Affect were the ones that were causally determined—does not, obviously at least, map on to Nietzsche's usage.

[15] Notice that the fact that conscious mental causes are not *adequate* to account for morally significant actions does *not* mean they are irrelevant to the best account. What is at stake in the claim about their lack of adequacy is only the thought that they are *type*-epiphenomenal, that is,

unconscious [*unbewusst*] processes from the accounting and to reflect on the preparation for an act only to the extent that it is conscious" (D 129), a view which Nietzsche plainly regards as mistaken, both here and in the passage quoted above. Indeed, the theme of the "ridiculous overestimation and misunderstanding of consciousness" (GS 11) is a recurring one in Nietzsche. "[B]y far the greatest part of our spirit's activity," says Nietzsche, "remains unconscious and unfelt" (GS 333; cf. GS 354). To be sure, there is a somewhat suspect overreaching in these passages: are we to believe that no conscious belief is part of the causal explanation of *any* action? I do not see that Nietzsche needs to defend this radical thesis, for what he is interested in debunking is the causal nexus between the conscious experience of will and actions of moral significance, that is, the actions for which moral praise and blame might be ascribed.[16]

Notice, of course, that there is, as of yet, no *argument* given for these claims, just conclusory statements about the causal inertness of putative conscious mental causes. But, once again, if we accept the Doctrine of Types underlying

their causal relevance is dependent upon the relationship in which they stand to physiological or unconscious causes. (Thanks to an anonymous referee for pressing this point.)

[16] Gemes (this volume, esp. at note 8) claims that Nietzsche was not mainly interested in the connection between free will and moral responsibility, but bases this on the quite surprising (and undocumented) assertion that Kant was not centrally interested in this issue. Gemes introduces a doctrine he calls "agency free will" (which he says certain later German Idealists—with whom, of course, Nietzsche had little familiarity or interest—were primarily interested in) which, he says, is more interested in "autonomy" than "moral responsibility." "Autonomy" on Gemes's view answers the question "what constitutes an action as opposed to a mere doing." To be sure, any view about free will and moral responsibility will have to distinguish between "doings" and "actions," but what Gemes owes us is both (1) an account of "actions" which somehow prescinds from claims about responsibility, and (2) evidence that this was Nietzsche's primary concern. As to (2), the best Gemes can do is call attention to the beginning of the Second Essay of the *Genealogy*, where Nietzsche writes about "the sovereign individual . . . with his own independent long will . . . This being who has become free, who is really *permitted* to promise, this lord of the *free* will, this sovereign . . ." (GM II 2). Even putting to one side the idiosyncratic nature of this passage—it is hard to see why this one passage should be thought more significant than the extensive textual evidence considered in the body of this paper—its meaning is far less evident than Gemes thinks. For Nietzsche quickly puts " 'free' " in quotes, to signal, of course, that this "sovereign individual" is not free in the sense requisite for moral responsibility (which is nowhere mentioned in the passage). And as I wrote about this section previously (Leiter 2002: 228), the sovereign individual "is distinguished by one overriding trait: he has "*the right to make a promise*" (GM II 2) because he can actually pull it off, i.e., his behavior is sufficiently regular and predictable so that he can be "answerable for his own future" (GM II 1), and he is able to remember what he has promised and honor that memory. But we know already from the opening lines of the Second Essay that this just means he is a certain kind of "animal" that has been bred the right way and nothing more. Indeed, should we be misled by the rhetoric of GM II 2, Nietzsche soon reminds us of the real point when he asks in the very next section: "How do you give a memory to the *animal* man?" (GM II 3, emphasis added)". The answer to that latter question about *breeding* involves a whole series of mnemonic techniques (many involving torture of one kind or another) that would, indeed, explain how an animal could be trained that is able to "keep a promise." Yet far more is at stake in the question of free will and moral responsibility than breeding animals who can make and discharge their promises, and nothing in the *Genealogy* suggests otherwise. (I should add that I am largely sympathetic to Gemes's thesis that Nietzsche views the *self* as a hierarchy of drives, but am unclear why he thinks Nietzsche conflates this point with the theory of the will I reconstruct in this essay.)

the first error, then the claims here also make some sense. That is, if morally significant actions and the conscious mental states that precede them are themselves the product of type-facts, then it follows that *the conscious mental states* that precede the action and whose propositional contents would make them appear to be causally connected to the action are, in fact, epiphenomenal, either as *tokens* or as *types*: that is, they are either causally inert with respect to the action *or* causally effective *only in virtue of* other type-facts about the person.

That brings us to the final error about causation, what Nietzsche calls the error of "imaginary causes" (TI VI 4). This error occurs when we invent, *post hoc*, causes to explain certain phenomena in our experience, phenomena that are, in reality, the cause of our invention. Nietzsche uses the striking example of dreams, though I'll modify the content of his example. Suppose while dreaming in the early morning hours, a police car, with siren wailing, passes by my window, but does not wake me. More often than not, in to my dream will emerge a narrative which *explains* the sound: perhaps in my dream I suddenly find myself being pursued by police, in their cars, with *their* sirens wailing. As Nietzsche puts it: "The representations which were *produced* by a certain state have been misunderstood as its causes" (TI VI 4). That is, the dream police car, and the dream siren—themselves *actually* the product of the real, external sound of a siren—are now, in the dream, treated as the *causes* of that sound.

Once again, what is significant for Nietzsche is that, as he puts it, "*the whole realm of morality and religion belongs under this concept of imaginary causes*" (TI VI 6). Let us take just one of Nietzsche's examples. Christians, he says, might "explain" "agreeable general feelings" as being produced by "faith, charity, and hope . . . the Christian virtues" (TI VI 6). One feels well, at peace, content, because one practices these Christian virtues—or so the religious explanation goes. But, objects Nietzsche,

> [A]ll these supposed explanations are resultant states and, as it were, translations of pleasurable or unpleasurable feelings into a false dialect: one is in a state of hope *because* the basic physiological feeling is once again strong and rich . . . . (TI VI 6)

So the Christian says, "That you have practiced the Christian virtues explains why you feel well and are at peace with yourself." In fact, says Nietzsche, there is a physiological explanation for why an agent who feels at peace with himself feels that way, and it is also an explanation for why he practices hope, faith, and charity. The structure of the criticism suggests, in fact, that the "error of imaginary causes" is just an instance of the first error, that "of confusing cause and effect," since, once again, one mistakes an "effect" (e.g., the feeling of hope) for the cause of something else (e.g., being at peace with oneself), when both are effects of an unrecognized "deep cause," i.e., "the basic physiological feeling" as

Nietzsche has it in this example.[17] And as with the earlier error, this one seems to depend entirely on accepting Nietzsche's Doctrine of Types, his doctrine that the psycho–physical facts about a person explain their conscious experience and behavior.

How does it follow from the three errors about causation—really, *two* errors, since the last is just an instance of the first—that "free will" is also an error? The error of confusing cause and effect is a *general* error that afflicts morality because morality is based on a mistaken picture of agency: we think that certain moral prescriptions will bring about certain consequences for those who follow them, yet the ability and disposition to act on the prescriptions, and the enjoyment of the consequences are possible only for certain *types* of persons. An exercise of free will plays no role. Notice that this argument makes no claim about the phenomenology of willing.

The error of false causality is an error because we wrongly infer that we know what causation is from our *experience* of the will being causal; but the will is not, in fact, causal, which follows from the Doctrine of Types. But, on any account of free will and moral responsibility, the will must be causal (even if not *causa sui*), in order for agents to have free will and be morally responsible for their actions. Therefore, if the error of false causality is a genuine error, then it follows that there is no free will. Only this second error implicates the phenomenology of willing, since it claims that we are in error in thinking we know what causation is based on our *experience* of the will. And the argument says we are in error here because our experience of the will misleads us as to the causal powers of the will: "there are no mental causes at all" Nietzsche tells us.

Let us return, now, to the question with which we started: why, according to Nietzsche, is the phenomenology of willing not a reliable guide to the causation of action or, more precisely, the causation of action such that we are morally responsible for the actions so caused? Nietzsche now has given us two arguments for the conclusion: the argument from the phenomenology of thoughts; and the argument from false causality, which depends on his Doctrine of Types.

THE REAL GENESIS OF ACTION

If the *experience* of willing does not, according to Nietzsche, illuminate how actions are brought about, what, then, *really* explains our actions? The "Four Great Errors" section of *Twilight of the Idols*, and the Doctrine of Types on

[17] So, too, in the case of dreams: one treats the dream police car as cause of the dream siren, when in fact both are caused by the *real* siren. What makes the case of "imaginary causes" a special instance of the error of confusing cause and effect is that in this case, E2 is itself a reflection of the deep cause. That difference, as far as I can see, does not matter for our purposes in the text.

which it depends, suggest an answer. *Type-facts*—facts about the unconscious psychology and the physiology of agents—explain our actions. That, alas, is to paint with a broom, not a brush, and does not yet shed much light on the actual causal structure of agency or its connection to the phenomenology of willing painted in such detail by Nietzsche in the section from *Beyond Good and Evil* we discussed at the start.

Here is where it may help to turn to some recent work in empirical psychology, powerfully synthesized by Daniel Wegner in his recent book *The Illusion of Conscious Will* (Wegner 2002). Wegner, like Nietzsche, starts from the *experience* of willing, and, like Nietzsche, wants to undermine our confidence that the experience accurately tracks the causal reality. To do so, Wegner calls our attention to cases where the *phenomenology* and the *causation* admittedly come apart: one set of cases involve "illusions of control," that is, "instances in which people have the feeling they are doing something when they actually are not doing anything" (2002: 9) (think of a video game, in which you feel your manipulation of the joy stick explains the action on the screen, when in fact, the machine is just running a pre-set program). Another set of well-documented cases involve the "automatisms," that is cases where there is action but no "experience of will" (2002: 8–9) (examples would include ouija board manipulation and behaviors under hypnotism). Wegner remarks:

> [T]he automatisms and illusions of control . . . remind us that action and the feeling of doing are not locked together inevitably. They come apart often enough to make one wonder whether they may be produced by separate systems in the mind. The processes of mind that produce the experience of will may be quite distinct from the processes of mind that produce the action itself. (2002: 11)

If the cases in question do, indeed, show that phenomenology of willing is not always an accurate guide to causation, they certainly do not show that this is *generally* true. But Wegner wants to establish Nietzsche's claim, namely, that the phenomenology of willing systematically misleads us as to the causation of our actions. And in the place of the "illusion of free will" as Wegner calls it, he proposes a different model according to which "both conscious willing and action are the effects of a common unconscious cause" (Holton 2004: 219), but the chain of causation does not run between the experience of willing and the action; rather, in Nietzschean terms, some type-fact about persons explains both the experience *and* the action (see the helpful diagram in Wegner and Wheatley 1999: 483; Wegner 2002: 68). As Wegner sums up his alternative picture of the causal genesis of action:

> [U]nconscious and inscrutable mechanisms create both conscious thought about action and the action, and also produce the sense of will we experience by perceiving the thought as cause of the action. So, while our thoughts may have deep, important, and unconscious causal connections to our actions, the experience of conscious will arises

from a process that interprets these connections, not from the connections themselves. (Wegner 2002: 98)

Before we turn to one part of Wegner's empirical evidence for this strong claim, it is important to recall an ambiguity that emerged in our earlier discussion of Nietzsche's epiphenomenalism, an ambiguity well-illustrated in the causal story suggested by Nietzsche's example of Cornaro. On one reading—call this *the Will as Secondary Cause*—slow metabolism (the relevant type-fact about Cornaro) explains why Cornaro ate a slender diet, and the fact that he ate a slender diet explains his longevity. If we take this version as an analogue of willing, then the will is, indeed, causal, but it is not the *ultimate* cause of an action: something *causes* the experience of willing and then the will causes the action.

On another reading—*the Will as Epiphenomenal*—the slow metabolism explains *both* the slender diet and the longevity, but there is no causal link between the latter two. The Cornaro example itself most plausibly suggests the Will as Secondary Cause (surely the slender diet makes a *causal* contribution to the long life), yet the passage with which we began, from *Daybreak*, suggests the Will as Epiphenomenal instead: if the "I will" is really analogous to the person "who steps out of his room at the moment when the sun steps out of its room, and then says: '*I will* that the sun shall rise' " (D 124), then there is no causal link between the experience of willing and the resulting action, just as there is no causal link between the person who wills the sun to rise and the rising of the sun.

Nietzsche's texts on this subject are, I believe, generally ambiguous as to which view of the will he decisively embraces. Wegner's empirical evidence, by contrast, is offered in support of the Will as Epiphenomenal. If Wegner is right, that is good reason, as a matter of interpretive charity, to read Nietzsche as committed to the latter view,[18] that is, to read him as holding the view that is (a) supported by his texts, and (b) most likely to be correct as a matter of empirical science.[19]

Wegner adduces a variety of kinds of evidence to support the idea of the Will as Epiphenomenal, and it will not be possible to review most of them here. I want to concentrate on one illustrative bit, already well known to students of the free will literature, but probably less familiar to those interested in Nietzsche. These are the studies by Benjamin Libet and colleagues (discussed in Wegner 2002: 50–5) examining the brain electrical activity (the "readiness potential" or "RP") that precedes an action (such as moving a finger) and the experience

[18] In the terms used in Leiter (2002: 91–2), this would be to read Nietzsche as embracing Token Epiphenomenalism, contrary to the interpretation emphasized in my earlier work.

[19] To the extent the Will as Epiphenomenal is not vindicated by empirical research—the verdict is plainly out—then the argument from interpretive charity may ultimately cut the other way. (Thanks to Eddy Nahmias for pressing this point.)

of willing. What the researchers found is that "the conscious willing of finger movement occurred at a significant interval *after* the onset of the RP (and also at a significant interval before the awareness of the movement)" (Wegner 2002: 53). According to Wegner, "These findings suggest that the brain starts doing something first . . . .[t]hen the person becomes conscious of wanting to do the action" that the brain has already initiated (Wegner 2002: 53).[20] Wegner (2002: 54) quotes Libet summing up the import of his findings as follows:

> [T]he initiation of the voluntary act appears to be an unconscious cerebral process. Clearly, free will or free choice of whether *to act now* could not be the initiating agent, contrary to one widely held view. This is of course also contrary to each individual's own introspective feeling that he/she consciously initiates such voluntary acts; this provides an important empirical example of the possibility that the subjective experience of a mental causality need not necessarily reflect the actual causative relationship between mental and brain events.[21]

In other words, about a century after Nietzsche, empirical psychologists have adduced evidence supporting his theory that the phenomenology of willing misleads as to the actual causal genesis of action.

This is not the only place that empirical psychology has "caught up," as it were, with Nietzsche—or, more charitably, provided confirmation for Nietzsche's theory of the will. Wegner adduces support, for example, for what Nietzsche calls the "error of false causality" (Wegner 2002: 64 ff.), as well as the "error of confusing cause and effect" (Wegner 2002: 66 ff., 96 ff.). What is worth emphasizing here, however, is the extent to which the alternative picture of the genesis of action suggested by Libet and Wegner—namely, that unconscious processes explain both the action and the experience of willing—resonates with Nietzsche's own, more detailed remarks on the genesis of action that come in the fascinating passage in Section 109 of *Daybreak* on "self-mastery."

Nietzsche begins this discussion by canvassing six different ways of "combating the vehemence of a drive [*eine Triebe*]." What follows is Nietzsche at his most

[20] This experimental result has a rather striking resonance with a phenomenological claim made by Nietzsche in *Nachlass* material, namely, that "in cases of sudden pain the reflex comes noticeably earlier than the [conscious] sensation of pain" (WP 699). (Thanks to an anonymous referee for calling this passage to my attention.)

[21] Alfred Mele (2006: 30–48) has recently mounted a sustained attack on Libet's interpretation of his results. Mele says that the onset of the conscious experience of willing after the RP "leaves it open that . . . rather than acquiring an intention or making a decision of which he is not conscious [at the time of the RP], the agent instead acquires an *urge* or *desire* of which he is not conscious," and which only becomes effective in virtue of the conscious intention (2006: 33). Although Mele's point gains some support from an interpretation of other experiments by Libet (discussed by Mele at 34–40), he establishes only that there is an *alternative* interpretation, not that his alternative is correct. But the resolution of the Libet–Mele dispute does not matter for Nietzsche's purposes, since Nietzsche is an incompatibilist, and the Libet results, even on Mele's rendering, show that *the causal trajectory* (whether that is an urge or an intention) leading to the action begins prior to the conscious intention to perform the action (whenever that occurs), and that is sufficient to defeat the *causa sui* conception of freedom of the will.

psychologically astute, as he documents six different ways of mastering a powerful urge: for example, by avoiding opportunities for gratification of the drive, thus weakening it over time; or by learning to associate painful thoughts with the drive, so that its satisfaction no longer has a positive valence.

Interesting as these observations are, the real significance of this passage for our purposes lies elsewhere. For Nietzsche is also concerned to answer the question as to the "ultimate motive" for "self-mastery." He explains it as follows:

> [*T]hat* one *wants* to combat the vehemence of a drive at all, however, does not stand within our own power; nor does the choice of any particular method; nor does the success or failure of this method. What is clearly the case is that in this entire procedure our intellect is only the blind instrument of *another drive*, which is a *rival* of the drive whose vehemence is tormenting us. . . . .While "we" believe we are complaining about the vehemence of a drive, at bottom it is one drive *which is complaining about the other*; that is to say: for us to become aware that we are suffering from the *vehemence* of a drive presupposes the existence of another equally vehement or even more vehement drive, and that a *struggle* is in prospect in which our intellect is going to have to take sides.

Even if the intellect must "take sides" (*Partei nehmen*) this plainly does not mean that the intellect determines which side prevails: to the contrary, the intellect is a mere spectator upon the struggle. Thus, the fact that one masters oneself is *not* a product of "free will," but rather an effect of the underlying type-facts characteristic of that person: namely, which of his various drives happens to be strongest. There is, as it were, no "self" in "self-mastery": that is, no conscious "self" who contributes anything to the process. "Self-mastery" is merely an effect of the interplay of certain unconscious drives, drives over which the conscious self exercises no control. A "person" is an arena in which the struggle of drives (type-facts) is played out; how they play out determines what he believes, what he values, what he becomes. But, *qua* conscious self or "agent", the person takes no active part in the process. As Nietzsche puts the same point, later, in *Beyond Good and Evil*: "The will to overcome an affect is, in the end, itself only the will of another, or several other, affects" (117; cf. also GM III 17). The will, in other words, or the experience of willing (in self-mastery), is itself the product of various unconscious drives or affects. Which is, in slightly different terms, exactly the theory of the will that some empirical psychologists have arrived at one hundred years after Nietzsche.[22]

[22] This paper has been in circulation so long that there is already a secondary literature replying to it. I have refrained from making too many changes responding to published (or circulating) critiques, but hope to do so at a later date. I am grateful to the graduate students in my fall 2003 seminar on "Nietzsche and Ethics" at the University of Texas at Austin for their help in thinking about these issues; to the audience at my keynote address on this topic at the meeting of the Friedrich Nietzsche Society at the University of Sussex in September 2004; to the discussants at the conference on "Nietzsche and Naturalism" organized by Mathias Risse at the Radcliffe Institute at Harvard University in November 2004; to audiences at Cambridge University, the University of Manchester, Washington University in St. Louis, Georgia State University, and the University

REFERENCES

WRITINGS BY NIETZSCHE

Beyond Good and Evil, trans. Walter Kaufmannn (New York: Vintage, 1966).
Daybreak, ed. Maudemarie Clark and Brian Leiter, trans. R. J. Hollingdale (Cambridge: Cambridge University Press, 1997).
The Gay Science, trans. Walter Kaufmann (New York: Vintage, 1974).
On the Genealogy of Morals, in *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage, 1969).
Twilight of the Idols, trans. R. J. Hollingdale, in *Twilight of the Idols and The Antichrist*, trans. R. J. Hollingdale (Middlesex: Penguin, 1968).
The Will to Power, ed. Walter Kaufmann, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage Books, 1968).

OTHER WORKS CITED

Holton, Richard (2004). 'Review of Wegner (2002)', *Mind* 113: 218–21.
Gemes, Ken (2009). 'Nietzsche on Free Will, Autonomy and the Sovereign Individual' (this volume).
Katsafanas, Paul (2005). 'Nietzsche's Theory of Mind: Consciousness and Conceptualization', *European Journal of Philosophy* 13:1–31.
Leiter, Brian (1998). 'The Paradox of Fatalism and Self-Creation in Nietzsche', in Christopher Janaway (ed.), *Willing and Nothingness: Schopenhauer as Nietzsche's Educator* (Oxford: Oxford University Press).
____ (2002). *Nietzsche on Morality* (London: Routledge).
Mele, Alfred (2006). *Free Will and Luck* (Oxford: Oxford University Press).
Nagel, Thomas (1986). *The View from Nowhere* (New York: Oxford University Press).
Owen, David and Ridley, Aaron (2003). 'On Fate', *International Studies in Philosophy* 35/3: 63–78.
Pereboom, Derk (2001). *Living Without Free Will* (Cambridge: Cambridge University Press).
Strawson, Galen (1994). 'The Impossibility of Moral Responsibility', *Philosophical Studies* 75: 5–24.
Wegner, Daniel M. (2002). *The Illusion of Conscious Will* (Cambridge, Mass.: Harvard University Press).
____ and Thalia Wheatley (1999). 'Apparent Mental Causation: Sources of the Experience of Will', *American Psychologist* 54: 480–92.

of Minnesota, Minneapolis-St. Paul; and to Reid Blackman, Daniel Came, Richard Holton, Paul Katsafanas, Derk Pereboom, and two anonymous referees for detailed comments.

7

Nietzsche's Freedoms*

John Richardson

It's well known that Nietzsche expresses both positive and negative views about freedom. Both, indeed, turn important wheels in his philosophy. His denials of the possibility or coherence of free will in the traditional sense—as a first cause, an uncaused cause—are often-stressed arguments in his attack on morality, which he claims relies on the picture of us as free in this way. But my focus will be on his positive notion—on the different kind of freedom he asserts we can have. I think this positive idea of freedom is even more important in his thinking, since it is one main way he describes and justifies his ideal to us.[1]

Now these positive and negative views may be taken as inconsistent. It seems they must be, if these views are about 'the same thing'.[2] For Nietzsche often expresses his negative verdicts on freedom as unqualified denials: freedom or free will is impossible, the very idea is incoherent. These denials are generally not tempered as against freedom 'in one sense'. Moreover Nietzsche similarly rejects as nonexistent or impossible many of the other things we closely associate with freedom, including self, will, and responsibility. So he seems to lay waste the whole ground around freedom, leaving nothing with which to build a positive account.

Nevertheless I think he has such an account, and that we only understand those denials by subordinating them to the positive notion. So every one, I will argue, of these assertions that 'there is no X'—each outright denial of some usual term for agency and personhood—should be heard as a provocative and

* Translations from Nietzsche's work in this paper are either mine or are my modifications of published translations. I have used the standard abbreviations and section numbers for purposes of reference.

[1] My second main aim, as will emerge, is to develop Nietzsche's account—and critique—of *agency*, in a way that goes beyond my treatment of it in (1996) and (2004).

[2] Might Nietzsche affirm freedom *in the same sense* that he denies it? Some possibilities here: a) he changes his view (in some decided way); b) he has no stable view, shifting freely among inconsistent ways of viewing it; c) he says it not about freedom in different senses, but speaks about the same freedom as it were in different voices, e.g. sometimes theoretically sometimes practically. I'll discuss the third of these below. Regarding the first, see Janaway (2007), arguing that in *Human, All Too Human* Nietzsche's view is more simply negative. I focus my attention on the later works, and above all on the *Genealogy*.

simplifying statement of a nuanced point about what X really is. In every case he wants to speak about Xs himself—to make claims that depend on there being such things. So in particular for freedom: he tells us what it's not, as a step to the really important point, what it is.

Of course the things Nietzsche most commonly insists to be unreal or impossible are things (like freedom) that have always been valued. He singles out his targets with this eye for what we care about, so that his denials are cutting, not reassuring. He means to disturb us that we lack the freedom we've presumed and been proud of. He also tries to disturb this value another way, which is even more vital. For he has another kind of negative point in reserve behind those denials of existence: he concedes that these Xs can and do occur, but argues that they're a bad thing. So this freedom we've presumed, insofar as we really can have it, is in fact nothing to be proud of, but something damaging and diminishing. The way we have been free is a shabby and narrowing thing.

And yet, even this deeper valuative attack is a part of a larger, positive view—in which Nietzsche takes up freedom, and being a self, and even responsibility, and makes them into a fresh ideal of his own. Far from freedom being either impossible or damaging, it's the way he most calls us to be. Indeed I think even the casual reader takes Nietzsche as promoting and exemplifying a striking new freedom, even though there are few explicit announcements, and no extended developments of the point.

One conspicuous[3] place he affirms and praises freedom this way is in GM II 2, where he speaks of 'the *sovereign individual* [*souveraine Individuum*], who resembles only himself, who has come loose [*losgekommene*] again from the ethic of custom, autonomous [*autonome*] and supraethical [*übersittliche*] (for "autonomous" and "ethical" are mutually exclusive), in short, the human being with his own independent [*unabhängigen*] long will', with 'an authentic consciousness of power and freedom, a feeling of the completion of the human in general' (GM II 2).[4] We'll see that there are difficulties in this passage, but it is a dramatic statement of his positive view.

So the challenge is to say what this freedom is, in the sense in which Nietzsche affirms and preaches it, and to distinguish this from the sense in which he denies and rejects it. In doing so we will need to situate freedom with respect to other key Nietzschean concepts in its vicinity. These include power, self, and responsibility. Let me say some first things about these connections, all three of which are evident in GM II 2.

[3] Recently much discussed; I have learned from earlier versions of all the other papers in this volume.

[4] Here 'ethic of custom' translates *Sittlichkeit der Sitte*; since Nietzsche means by this a way of valuing of social rules and conventions, prior to morality and still operative 'beneath' it, I think it's better not to use 'moral' in rendering *sittlich*. Preferable is 'ethical', especially because we can hear 'ethos' strongly present in it.

Power is Nietzsche's ur-value, so to understand why he values freedom we must see how it is related to power. That he closely connects these is evident in the quote from GM II 2 above, which attributes to the sovereign individual a consciousness of power and freedom; Nietzsche soon adds: 'mastery over himself also necessarily gives in hand mastery over circumstances, over nature and all shorter-willed and more unreliable creatures'. More succinctly, WP 770: 'freedom understood, that is, as positive power, as will to power'.[5]

Becoming free is also very intimately related to becoming a *self* or agent. Familiarly, Nietzsche thinks that a self is something that needs to be acquired, or rather created—this is one main reason for calling him an existentialist. Only a self can be free, and freedom may be no or little more than most fully being a self. These goals lie in the same direction, are points along the same axis for Nietzsche (as for other existentialists). We find their connection displayed in GM II 2's hero-name 'the sovereign individual'—his achievement is both to be free and to be an individual.[6]

Traditionally, one of freedom's main roles has been to ground *responsibility*: we can only be responsible, insofar as we are free. We might expect Nietzsche to be a thorough critic of responsibility, as a kind of guiltiness, and so to mean to sever this tie, even if he retains and revalues freedom.[7] But in fact he keeps even this connection. The freedom he advocates involves a magnified sense of responsibility, as we likewise see in GM II 2, which attributes to the sovereign individual 'the proud knowledge of the extraordinary privilege of *responsibility*, the consciousness of this rare freedom, this power over oneself and fate'.

I will try to show how Nietzsche offers a new conception of freedom, which values it as a privileged kind of power, achieving a new kind of self, and assuming a new kind of responsibility. So, generally, his critique is of past or existing notions and versions of freedom, and his positive views are about what it can be. The former is his theory of freedom, meant naturalistically; the latter is a projective valuing of what freedom should (and can) be. So he carries out a 'redesign' of freedom and these allied notions. This redesign has, I think, two principal objectives.

Nietzsche wants in the first place to '*naturalize*' freedom—to show it as a feature of a certain kind of organism, one variety of 'life'.[8] He traces the roots

[5] Also GS 347, which begins by saying that the extent to which someone needs 'faith' to flourish is 'a measure of the degree of his strength', and concludes that 'a delight and force of self-determination, a *freedom* of the will is thinkable, in which a spirit takes leave of every faith [*Glauben*], of every wish for certainty'. GM II 18 refers to 'that *instinct for freedom* (spoken in my language: the will to power)'.

[6] Katsafanas (unpub., ch. 6) presents a strong case for distinguishing selfhood from freedom in Nietzsche, with which I partly concur. But I develop a rather different account of their relationship below.

[7] See how TI VI 8 says that 'the great liberation' depends on nobody being held responsible any more.

[8] In this, his account is comparable to Dennett's well-known one in *Freedom Evolves*.

of freedom back to our animality; the latter is still operative in it. The freedom he rejects as impossible is the freedom of a metaphysical subject—somehow detached from that biological root, something 'pure'. It is therefore crucial that freedom is something that has evolved, within a certain kind of animal, from abilities that were at first merely those of other animals—and that freedom is something that is evolving still.

This naturalizing story about freedom will also, necessarily, be about 'freedom', i.e. an idea or notion of freedom. As freedom has evolved in this (kind of) organism, it has done so in tandem with a conception this organism has had of itself *as free*. And Nietzsche is as much interested in this idea—and aim and value—of freedom as he is in freedom itself. This point greatly complicates the logic of his discussion.

Freedom, we'll see, is a certain skill or capacity, a *dunamis*, passed on from organism to organism as an instinct or habit. This capacity has evolved, by biological and by social processes, through different forms in different historical settings. And in each case the skill is associated with a certain view of itself, an idea of what is being done, of what this freedom is it's achieving. The skill itself involves a certain perspective, most importantly the perspective on itself, of what it is and is trying to be.[9] Freedom is a skill aimed at an idea of freedom. We need to track the evolution of skill and idea in tandem. One crucial question will be to what extent the skill's perspective on itself is true. As we practice freedom, do we conceive ourselves to be free in the very way we are?

So Nietzsche tells an historical story about freedom and its concept.[10] We must therefore apply to them the general lesson he draws about historical phenomena, in the famous treatment of punishment in GM II 13. Freedom too is indefinable, because the concept and practice both bear the diverse meanings that have been layered into them by that design history. The word bears that complexity, in its many uses in different contexts, and the practice does too, in its diverse parts shaped and settled at different times (under different selective forces) in the past. Some of the difficulty in Nietzsche's treatment of freedom is that he speaks readily within the many different senses thus built into the term.[11]

I said that these past meanings are 'layered into' later ones. This is a very important point for Nietzsche: the cumulative character of this history is the

[9] GS 353 says that founders of religions invent both a) a practice and b) an interpretation that gives this practice highest value.

[10] Nietzsche is already telling this story in D 18: 'Every smallest step in the field of free thought, of a life shaped personally, has always had to be fought for with spiritual and bodily tortures. . . . Nothing has been purchased more dearly than this little bit of human reason and feeling of freedom that constitutes our pride.'

[11] That is, he uses 'freedom' not as a settled and technical term, but speaks with the same multiplicity that has been embedded in the term by its history—with an additional sense of his own (and which supplements rather than replaces the others). So notice how he takes over the word's special senses in contexts: artistic freedom (BGE 188, 213), freedom as opposed to servitude or slavery (GM II 5), freedom as 'letting go' (TI IX 41), and many others.

reason historical things are indefinable. And we'll see that freedom, as an evolving ability, incorporates and builds on its earlier stages, so that the simpler kinds of freedom are components of the later ones. However it should also be emphasized that this preservation of the past is only partial.[12] In the case of punishment, for example, there may well be past uses that have been so completely overwritten that no trace remains of them in our current practice and idea of punishment. Nietzsche's hope is to overwrite and expunge some of what freedom has been.

On this view there is a sense in which freedom is 'one thing' after all—it is in fact an extended, historical thing, an evolving ability with an idea of itself. So Nietzsche is saved from self-contradiction, in his various remarks about freedom, not by speaking of different things (different kinds of freedom), but by talking about different parts of this one historical thing. The seeming contradictions are resolved when we see that he describes different temporal parts of this evolving thing, and treats sometimes its ability, sometimes its idea of itself. They are resolved when we view them with the 'historical sense' he laments philosophy has generally lacked (TI III 1).

It's important that this development is still going on. Nietzsche in fact means to participate in it—to help us to carry out a further redesign, appropriate for us today. In his most favored and enthusiastic and whole-hearted uses of the term, he is not analyzing what freedom has been or is, but advocating what it should be, promoting his redesign. So we should distinguish his naturalistic analysis of the concept and practices so far, and his 'revaluation', by which he promotes a new phase in the evolution of freedom. (The new freedom will turn out to crucially include this process of genealogy and revaluation, itself.)

These last points have brought us to the second principal objective Nietzsche has in revising 'freedom': he wants crucially to *de-moralize* it. This is the main character of his 'revaluation of values'. He revalues freedom not in the sense of changing a positive evaluation into a negative, but by redesigning it to serve different ends. He takes it out of service of moral ends, and adapts it to serve the new end—which he thinks is life's old end—of power.

The point is to change what freedom is for, what outcomes it tends to bring about. So far, as we'll see, this practice (-and-concept) has been designed to socialize us, to make us better herd animals. In recent millennia it has done so by making people moral (this connection is expressed revealingly in Kant). Nietzsche gives reasons for resisting and refusing this moral ideal. He promotes a personal and also a social project to his readers, to change the practice to serve ends he advocates in place of morality's.

In '*de-moralizing*' freedom Nietzsche's point is not, then, to render it valuatively neutral. The aim isn't to view freedom solely in a naturalistic or scientific spirit, to strip the practice or concept of all valuative implications whatsoever.

[12] GM II 12 says that 'the previous "meaning" and "purpose" must of necessity become obscured or entirely extinguished'.

Nietzsche still wants a practice of pursuing and desiring freedom, in which the concept counts as an ideal. Besides his naturalizing genealogy of freedom, which subjects freedom to cold and critical study, Nietzsche also speaks of the term as a value, modeling the new practice of freedom he promotes.

We might sum these points up with the slogan: freedom is not, it becomes. It must be studied as a long cultural process, within which we now discover ourselves to stand at a certain point. Freedom now means what its history has shaped it to mean—and the latter is what genealogy discovers. But freedom is still becoming, and we can participate in this, by carrying out a revaluation of it in the light of that genealogy.

In Nietzsche's story how freedom becomes there are—I will argue—three principal phases or stages.[13] Each of these is a certain capacity or skill, which involves, as we've seen, a certain idea of itself, of itself as 'free'. These ideas vary, we shall see, in how true or accurate they are as to the skill itself. Each skill is a kind of power, each constitutes a kind of self and assumes a kind of responsibility. These skills arise and evolve under different selective forces, and are thereby designed to served different ends. But these kinds of freedom are not just stages, because earlier skills 'accumulate' within later ones. So these are also the three elements in the full conception of freedom Nietzsche promotes to us. Here we meet, in turn, Nietzsche's notions of selfhood, agency, and (full, Nietzschean) freedom.

a) Unifying the drives. The first account of freedom is the most evident in Nietzsche's texts. Its main feature is to locate freedom—and the associated notion of an I or self—at the level of the drives, in terms of the drives. This move 'down' to the level of the drives in explaining thought and action is characteristic of Nietzsche—although we'll see (in b) that he also constitutes a kind of agency on top of this. This underlying psychology of drives has an almost chemical simplicity. Freedom of this first sort is a skill in and of these drives.

Here Nietzsche thinks of freedom as a capacity that reaches down into our animality, into our condition simply as (an instance of) 'life'. I think he would be willing to attribute this first kind of freedom to organisms generally, though of course he is always focused on the human case.[14] For each living thing is a collection of drives, drives that compete to control what the organism does. It's when he treats freedom just in terms of these drives that he most obviously 'naturalizes' it (though really the rest of his story will be naturalistic as well).

When Nietzsche denies that there is any I or self or subject, what he usually puts in its place is that collection of drives. For example D 119 speaks of 'the

[13] This tripartite schema is typical for Nietzsche—he might get it from Hegel, or from Plato, or from anywhere else. We find it in his schema of animal, human, and overman, or of master, slave, and overman—and the three kinds of freedom map onto these.

[14] We'll see that in us, as the 'sick animal', this freedom often needs to be *re*-achieved, so that the natural freedom looks like a special accomplishment.

totality of *drives* that constitute [a man's] being'. And KSA 10: 274: 'As cell stands physiologically beside cell, so drive beside drive. The most general picture of our essence is *an association of drives*, with continual rivalries and alliances with one another.' When I act, I think that I as subject or self have chosen to do what I'm doing. But—he often says—there's no such I, and what really explains is some drive that just then had control of the organism. Throughout Nietzsche's psychology, it is these drives that are his favorite and usual explainers. Nevertheless there is a kind of freedom that occurs even here—indeed he finds freedom here in several ways.

The first way Nietzsche interprets freedom in terms of drives is quite simple: freedom is a drive's own feeling of power, while it commands another drive. It is the 'affect of command' that he so often mentions. Here freedom seems attributable to any single drive, at the moment it is able to command. Its opposite is the state and feeling of being hemmed in and constrained, prevented from venting or discharging. Consider some passages: KSA 11: 606 (WLN 36): 'the will . . . is above all also an *affect*: that affect of command. What is called freedom of the will is essentially a feeling of superiority over the one who must obey: "I am free, he must obey"—this consciousness is present in every will . . .'. GS 347: 'will, as the affect of command, is the decisive mark of self-mastery [*Selbstherrlichkeit*] and force'. And KSA 11: 505–6 (WLN 16): 'Free: *means* "not pushed and shoved, without a *feeling of compulsion*" '; 'it is the *feeling of our preponderance of force* that we call "freedom of the will", the consciousness that our force *compels* in relation to a force which is compelled'.

This simplest capacity of freedom already has an idea of itself, and this idea is already partly false. For this commanding will looks only ahead, and doesn't see its own constrainedness. So it has the beginnings of the metaphysical faith in 'freedom of the will'. HAH I 18: 'We are hungry, but originally do not think that the organism wills to sustain itself, but that feeling seems to make itself valid *without cause or purpose*, it isolates itself and considers itself *voluntary* [*willkürlich*]. Thus: belief in freedom of the will is an original error of everything organic'.

However Nietzsche isn't content with this very simple idea of freedom. Elsewhere, more usually, he makes freedom a feature not of an individual drive, but of the organism as a collection of drives. He makes being-a-self and being-free consist in a certain unity or unification of the drives.[15]

What is it, first, that needs such unifying? I suggest that we picture the drives as a set of forces of various strengths, pushing for various goals. By the relation between their goals, they tend either to further or hinder one another. And by the relation between their strengths, they tend either to command or obey one another. So, we may say, by their vectors and strengths, they stand at any moment in a certain 'power-structure' with one another, a system

[15] Katsafanas (unpub., ch. 5) gives a sophisticated analysis of this unity.

of oppositions and alliances, forcings and compulsions. But further, we must imagine this structure as fluctuating rather drastically through time, as particular drives strengthen (perhaps stimulated by the situation) or weaken (perhaps when sated). Now what will unification be?

A preliminary kind of unity is already present in Nietzsche's fuller analysis of that 'affect of command'. For it turns out that this affect involves not just the feeling of command, but the simultaneous feeling of obeying. It requires that the organism be, and experience itself as, at once both commander and obeyer. Again KSA 11: 606 (WLN 36): 'A man who *wills*—commands a something in himself which obeys, or which he believes will obey. . . . [I]n a given case we are simultaneously the commanders and the obeyers, and as obeyers know the feelings of resisting . . .'.[16]

So my drives are not isolated viewpoints, but intermesh, and combine into a single opening upon the world, whereby the world looks and feels as it does to them all, at once. It is this synthetic viewpoint that has them together before it, and experiences at once both the commanding and the obeying drive, that possesses the full affect of command, a kind of proto-freedom. As I constrain one drive in order to exercise another, I experience the latter's command more completely, by feeling it too from the point of view of the part that obeys.

This first unity depends just on drives' coalescing into an overall viewpoint in which they are all elements.[17] And I think Nietzsche attributes this to organisms generally, inasmuch as all of them have multiple drives, simultaneously active, and their drives all involve perspectives. So every organism 'views' the world within various sets or projects towards it, and 'feels' their different prospects together. So it can experience itself as constraining one aim to suit another, as when a cat needs to eat in an uncomfortable posture, and is at once gratified and bothered.

But this is only a preliminary way Nietzsche thinks of drives as unified or united. He has something further in mind by the unity required for selfhood and freedom. For even a chaotic and akratic person could still 'share' drives' perspectives this way—could, while acting from a momentarily dominant drive, simultaneously feel other drives sacrificed to it. How further must drives come together, for there to be a self? I think it's hard to pin down Nietzsche's position here—there are several candidates.

i) Nietzsche might define this unity in terms of functionality or effectiveness. That is, a set of drives would be synthesized or unified insofar as they mesh effectively together, so that the organism 'as a whole' can work effectively on

[16] BGE 19: '"Freedom of the will"—that is the expression for the complex state of delight of the willer [*Wollenden*], who commands and at the same time counts himself one with the executor of the order'. Earlier: 'we . . . at the same time are the commanding *and* the obeying, and as the obeying we know the feelings of constraint, impulsion, pressure, resistance, motion'.

[17] Or perhaps not all of our drives are always active, and we see things at any moment only from the several stances that are active in us then.

its environment. I think Nietzsche does attribute this functional unity even to animals: when 'healthy', their drives are individually strong, but also synthesized, organized, and unified so as to maximize the organism's overall success. So particular drives operate when and to the extent that best serves the organism. Nietzsche often associates unity with such effectiveness. However I doubt that he identifies them; success is instead a contingent result. Unity has some character of its own, that makes it effective.

ii) So instead, unity might lie in the drives' coalescing into a *stable* power-structure—a *persisting* network of power-relations to one another. Each drive, that is, has a place in this structure, and is held to it. It doesn't sometimes overpower the other interests and run the organism for a while, to the detriment of the rest. So the organism keeps a consistent view of its interests, and runs its behavior with a steady aim. We can see how this kind of consistency might produce the effectiveness on the environment cited in i).[18]

I think that Nietzsche often does have this kind of unity in mind. He thinks of health as the capacity to achieve and sustain such a stable synthesis, making the organism more effective in turn. This capacity does not belong to some preexisting self, but to the collection of drives itself. A healthy organism is simply a set of drives that is able to settle into such a stable structure; it has a self, only by achieving that synthesis. Such health is usual in other animals, but not in humans, and especially in modern humans: for us this synthesis no longer 'comes naturally'. This is much of Nietzsche's view of unity, however not all.

iii) Unity may lie in there being *one* ruler of this set of drives—a single dominant drive, or perhaps a ruling 'committee' of drives. Although Nietzsche often has ii) in mind, he usually supplements it with this stronger point.[19] At least in the best case of unity, that synthesis of a stable power-system of drives is accomplished by a single drive taking control, and imposing its single command. WP 46: 'Multiplicity and disaggregation of the impulses, lack of system among them, results as "weak will"; their coordination under the dominance of a single one results as "strong will"—in the first case it is oscillation and the lack of a centre of gravity; in the latter precision and clarity of direction'.[20]

Unity, of one or more of these kinds, constitutes a first layer of agency, selfhood, and freedom. Most fully, this first freedom is the capacity to sustain a life under the command of a ruling passion. This capacity is free in both the positive and negative ways. Positively, it is the freedom to use the other drives,

[18] So perhaps Gemes (2009:46): 'to have a genuine self is to have an enduring co-ordinated hierarchy of drives'.

[19] Does unity in *animals* require this single ruling drive? Perhaps it's only humans, who have lost the natural drive-unity of ii), that need to reacquire it under that single rule.

[20] Nietzsche very often stresses the importance of a single dominant drive. WP 778: 'antagonism of passions . . .: very unhealthy, inner ruin, betraying and increasing an inner conflict and anarchism—unless one passion at last becomes master. *Return to health*'. Cf. Gemes (2009:42), who cites passages on Wagner's 'ruling passion' (UM III 2) and on 'the organizing "idea" that is destined to rule' (EH II 9).

and all the organism's resources, and thereby also things around it. Negatively it is freedom-from being controlled by isolated passions or impulses (and thereby also freedom-from control by the external stimuli of these passions or impulses).

This synthesis of the drives, under single command, matters especially because it is *power*—power over the other drives, i.e. over the host of the organism's capacities, and therefore power over nature. By unifying the drives it constitutes out of them a *self*. And this self is also *responsible*, in that it speaks for the organism, and can be relied on to direct the organism. The organism has its drives in control, and is as this synthetic whole responsible for its own behavior. As responsible, it is also reliable and (in this way) predictable to those around it.

The capacity for this unity belongs, we've seen, to our animality, but Nietzsche thinks this animal capacity has been damaged and suppressed. We face the challenge of re-attaining the ability to unify ourselves in this way. Individually, we need to hope that our system of drives has this capacity, and especially that we possess one drive that will become strong enough to rule.

In GM I Nietzsche presents a model for this kind of freedom in his 'masters'. Perhaps the master's key trait is precisely this capacity to organize himself. Slave morality's decisive blow to the master is then to disrupt this self-unity, to divide him against himself. In GM II, however, it emerges that the assault on our natural unity began much earlier, with the prehistoric taming of humans for social life, which already broke down the natural unity of our animal drives.

Nietzsche tells this first story about power and freedom completely in terms of the drives, which gives it a striking simplicity.[21] When he thinks along this line, he leaves little or no role for any conscious and deliberating self. Either there is no such thing, or it is an epiphenomenon (an aftereffect that affects nothing else), or it is simply a secondary tool or expression of the drives. And similarly, on this line, for our conscious values, articulated in our principles. These explain little, by contrast with a subconscious valuing that goes on in our drives.

b) The deliberative self. But this vocabulary of drives isn't enough for the explanations Nietzsche wants to give of us—nor for the new values he wants to offer us. In the end he means not to dismiss consciousness and deliberation, but to naturalize and de-moralize them. It's important to understand how we think and decide, and important to improve these, precisely because they are so important in how we live. Nietzsche criticizes the ways we are conscious and choose (the moral ways), and advocates different ways—neither of which would make sense if consciousness and choice were merely epiphenomenal. Moreover, it's clear that Nietzsche's works themselves appeal to these capacities in his readers—he wants

[21] BGE 117: 'the will to overcome an affect is ultimately only the will of another or several other affects'. D 109: 'the awareness of suffering from such *vehemence* [of one drive] presupposes that there is another equally vehement or even more vehement drive, and that a *struggle* is in prospect in which our intellect is going to have to take sides'.

us to reflect on his words, to deliberate in the light of them (even if perhaps he also means to affect us subliminally).[22]

So although, as I have said, drives are Nietzsche's principal explainers, they are not his sole explainers. How we reflect and choose are effective as well.[23] Indeed it's often through these that the drives themselves operate: Nietzsche very commonly points out surprising ways our decisions express unrecognized or unacknowledged drives. So even the story he tells about the drives often depends on an account of agency. And so we must not take those outright rejections of any such thing as conclusive.

Nietzsche not only needs reflection and choice as explainers, he also needs them in his values—they are features of the human conditions he values most. They are a further kind of freedom, built on the animal, and an element in Nietzsche's own. In a different way than the unity in the drives, reflection and choice involve the 'affect of command', and constitute kinds of power, self, and responsibility—kinds somewhat different from those we saw in a). This freedom also takes itself to be free, i.e. aims at an idea of its own agency or choice.

In naturalizing this '*agency*' (I'll use this term to sum up this broad topic) Nietzsche treats it, like the drives, as a certain capacity or ability of an individual organism. But whereas the capacity to unify the drives has roots back in our animal past, the capacity to decide was evolved in our human prehistory, and distinguishes us. The individual acquires this capacity not genetically but from the social matrix, where it has evolved through a long history, which Nietzsche offers various sketches of.

As a dispositional capacity, agency competes against the drives as other such dispositions. (Nietzsche thinks of them as competing to control the organism, by dictating its outlook and effort. Since the organism is really only the sum of its powers or dispositions, this competition is by each part to rule the whole, to organize the others into *its* whole.) Agency is reinforced or opposed by particular drives, as the latter are by one another. It is as it were 'in the trenches' with the drives, not something different in kind.

So agency is indeed a kind of drive itself, which I think is one reason Nietzsche sometimes denies that that there can be any self or agent. He thinks we make it essential to agency that it be something quite different in kind from the drives—and if this is essential, there can indeed be no agency. Moreover this conception of agency is not just an error of priests and philosophers. It belongs to the very 'idea' this capacity has of itself—what it takes itself to be trying to be. As we'll see, the very act of reflecting and choosing involves a

[22] The failure of agency to study itself is expressed in GM Preface: 'We are unknown to ourselves, we knowers'; the task of the book (and of his others) is to turn our attention back upon ourselves, as knowers.

[23] GS 335: 'our opinions, valuations, and tables of goods are certainly some of the most powerful levers in the machinery of our actions'. These are levers inasmuch as they are *especially* effective points through which to alter behavior.

self-misconception: even as it refers to itself as 'I' and 'self' and 'free', it gets itself wrong. All of this influences Nietzsche's denials of the I, self, and freedom.

Nietzsche offers various theories how this capacity for agency evolved—always along with that embedded idea of itself. He asks especially what forces shaped this skill, and to what purposes. He tries thereby to uncover what interests our capacity for agency expresses. I will focus on two of his discussions, on GS 354's account of the evolution of consciousness and language, and on the second essay of the *Genealogy*, which tells an elaborate story how the 'ability to promise' has developed. I suggest that we put these together as Nietzsche's account of the evolution of agency.

Let's look first at GS 354, which tells a story how and why consciousness and language arise. These are two key parts, I suggest, of the deliberative capacity we want to understand. The 'animal' drives are unconscious and nonlinguistic. How did these capacities evolve, and what functions were they selected to play?

Let me quickly present GS 354's story by quoting selectively from it. Nietzsche argues in this long paragraph that 'the development of language and the development of consciousness . . . go hand in hand', because both have the function to 'facilitate communication'. In the primary case, one becomes conscious of something about oneself in order to put it into words, and to share it with others. The drives were already well able to think, and what distinguishes conscious thinking is that it '*takes place in words*', hence is readied or prepared to be shared. So '[c]onsciousness is really just a binding-net [*Verbindungsnetz*] between person and person—only in this capacity did it have to develop; the solitary and predatory person would not have needed it.'[24] This means that 'consciousness actually belongs not to the individual-existence of a person, but rather to the community- and herd-nature in him'.[25]

Nietzsche infers (still in GS 354) that consciousness 'is finely developed only in relation to its usefulness to community or herd'. It is designed to serve the interest of the community in binding us more tightly up in it. And this is the effect it still has in us, so that 'each of us, even with the best will to *understand* himself as individually as possible, . . . will always bring to consciousness precisely the "not-individual" in him, his "averageness"—that through the character of consciousness . . . our thought itself is continually . . . translated back into the herd-perspective'.

Notice how Nietzsche's story runs against what we might have supposed: that consciousness and language were developed as tools for the drives—neutral tools, generally available to the drives and serving them. (We might have thought that by serving the drives they served the organism's interest in survival and,

[24] WP 524: consciousness 'is only a *means for communicability*'.

[25] GM II 16 likewise associates the development of an 'inner world' with humans' subjection to society, but as an effect: society prevents drives from discharging outwardly, so they 'turn themselves inwards'.

especially, power or growth.) On this model different drives at different times would take hold of these tools, so that it is now one drive that speaks and thinks, now another. But those tools—consciousness and language—would themselves be indifferent among them, since they evolved (on this supposal) to serve them generally.[26]

But Nietzsche here denies that language and consciousness are 'neutral' in this way—and that they indifferently serve the drives generally. These capacities themselves express interests, and so are analogous to drives themselves. Moreover the interests they serve are not the interests of the drives. They are those of society: these capacities are designed to serve a *socializing* function—to facilitate our 'taming' into good 'herd animals'. Rather than neutral servants of the drives, consciousness and language compete with them on behalf of those social interests; they represent those interests against the drives.[27] (Notice how strongly all of this counts against the idea of consciousness as a mere epiphenomenon. The problem here is that it's all-too-effective—at herding us.)

Let's turn now to GM II which describes the central capacity that consciousness and language both serve, the capacity central to our socialization. Nietzsche presents this as 'the ability to promise', but it is something more elaborate and far-reaching than this makes it seem. It's the broad phenomenon I'm calling agency.

Promising, first of all, is something complex: it depends on a number of interlocking sub-capacities. One of these is the capacity to remember what one has promised or committed to do. Another is the capacity to commit oneself (to remember in the future). This memory, further, is a 'memory of the will' (GM II 1), i.e. an effective memory that is able not just to retrieve commitments but to abide by them. That is, the capacity induces the whole organism to abide by them.

The promiser must be able to abide by commitments even or especially when strong drives incline it not to. So the capacity must include a strong *inhibitive* power, to refrain from acting immediately upon one's drives. The promiser is able to 'insert a pause' in which to consult its commitments, and determine what these require in this circumstance. In its strong power to restrain the drives, this capacity is 'ascetic'; here it draws, perhaps, on existing forces of restraint present independently in the drives. TI VIII 6: 'This is the first preschooling for spirituality: not to react immediately to a stimulus, but instead to take in hand the inhibiting, excluding instincts.'

For what purpose does this capacity restrain the drives? For the sake of one's commitments. GM II begins by talking about commitments to other individuals in relations of trade (see especially II 8), but Nietzsche soon shifts to what I

[26] But Nietzsche perhaps thinks this way in WP 524: 'Consciousness is *not* the director but an *organ of the director*'.

[27] Compare Katsafanas (unpub., ch. 5) on the 'vector model' of will, which makes it 'one motive among others'.

think is really more important to him: commitments to society. This capacity (for agency) is above all designed to give humans the capacity to remember (and abide by) the social rules. 'With the help of such images and processes [the horrific punishments inflicted on breakers] one finally retains in memory five, six "I will nots", in connection with which one has given one's *promise* in order to live within the advantages of society' (GM II 3; cf. also II 9).

What one commits to then, are these social norms. Or more broadly one commits to be normal, to be like others. This capacity (for promising, for agency) serves the overall function of socializing and taming us—and this is perhaps Nietzsche's most important lesson for us about it. This is the main interest this new capacity serves, and it serves it against the interests of the drives. It is the ability to refrain from acting upon drives, so as to refer to principles or rules (or to 'what one does'), to which one then conforms one's behavior. So although this new capacity works at the same level as the drives—we've seen that it is simply a dispositional capacity itself—it represents a kind of alien interest against them.[28] It is thus deeply 'ascetic'.

It's because this new disposition is so foreign and so contrary to the drives, that such hard training has been needed in order to breed it into us. It has had to be inserted amongst our animal drives and enabled to compete with them, indeed to control them. To simplify, we might say that there are two main phases of this training, for Nietzsche. In the earlier and much longer 'ethic of custom', the training worked mainly via individuals' fear—fear of the public torturing punishments. In the more recent phase, which Nietzsche calls 'morality', the training works more subtly by guilt, in which the individual's own aggressive drives are enlisted as enforcers of the moral rules: they take pleasure in punishing the part of the individual that goes astray.

We've seen that consciousness and language evolved with the function to socialize us. They do so in their own right: in learning a language the individual acquires a generic or average structuring, which we've seen constrains our conscious thinking as well. But consciousness and language also socialize us by supporting that ability to promise, i.e. to 'remember the rules'. Those rules are stated in language, and remembered as such. And the rules are applied in consciousness—they are made so as to work explicitly, by contrast with the purposes in the drives. (It's by having these conscious rules that the organism has 'values' in our usual sense—though Nietzsche uses the term more liberally, insisting that we, and other animals, already value in the unconscious drives.) As some drive inclines me towards an inviting act, memory must jolt me into awareness of the rule that prohibits it.[29]

[28] Cf. GS 296 on how society shapes us to be dependable, to have a reliable character in this way; it frowns on self-transformation.

[29] Conversely, we might say that learning a language itself involves promising: one undertakes a commitment to use words in the normal and common ways.

So Nietzsche's naturalizing story attributes to us—human organisms—a certain package of abilities. We might sum it as the capacity, in an 'action situation', to restrain our drives in order to become aware of certain relevant rules, and to act in accordance with these—hence, sometimes, against the interests of those drives. There are organisms with this capacity, through that long social training. But now where is the 'self' or 'I' in this story, and where is 'freedom'?

It is the organism that has this new dispositional ability, alongside the other abilities that are its drives. Yet there seems to us—doesn't there?—to be a self or I that has this new ability, but experiences the drives as external to it. 'I' am the subject of consciousness, it seems, which weighs alternatives and determines goals. 'I' carry out my conscious thinking and deciding, and for 'me' the influences of these drives and affects need to be uncovered and questioned: these drives don't speak for me, in the way that my deliberate choices do.[30]

On Nietzsche's naturalistic story, this notion of an I is the idea the new ability has of itself. The ability to consciously think and choose includes or involves a certain idea of itself—an idea that of course is itself conscious and linguistic. The capacity takes a first-person point of view: it calls and thinks itself 'I'. It also thinks various things about itself, many of which are false. For example it takes itself to be the essence or core of the organism, something different in kind from the drives, and so on.

It's here at last that we come to the naturalistic site of the idea of freedom as free will, as first cause, which we noted at the start is Nietzsche's principal target. This idea is a product of the second, moralizing phase of the long social taming of humans into agency. In GM I 13 Nietzsche depicts the slave as the particular vehicle of this innovation: the resentful slave needs the master to be free and responsible for his evil deeds, the better to punish him—or rather to imagine him punished. And on the other hand the slave wants to count his own weakness as really the result of his own choice to 'turn the other cheek'. Since the negative judgment is more vital to the slave, free will is needed especially for the purpose of grounding and intensifying blame.

But agency's idea of its own freedom and responsibility is not merely a product of this specific human type (the reactive slave). It has deeper roots, and a broader role. It is, ultimately, a further means to enhance social control of members, via their own agential control of their drives. Agency's principal function (what social processes have mainly designed it for) is to ensure that members do what they must for the society to hang together and prosper. It does so by, in particular,

[30] GS 305: 'Those moralists who command the person first and above all to take control in himself thereby bring a peculiar disease upon him: namely, a constant stimulability by all natural stirrings and inclinations . . . . Whatever may henceforth push, pull, beckon, impel him from within or without will always strike this stimulable one as endangering his self-mastery'.

making agents predictable to one another (able to promise), the better to mesh their activities (to trade). And it makes them predictable by making them able and willing to frustrate their drives, and—importantly—to bear up under the great suffering and physical depression Nietzsche thinks this caused.

Agency's idea of itself serves this function. It takes itself as able (free) not only to restrain particular drives—which it often is—but to choose and will in complete independence from them all. This illusion both enspirits it in its struggle against drives, and prepares it to feel guilty wherever it fails.[31] So the idea of free will is an ally of bad conscience, which is the distinctive *moral* form of social control through values. It harnesses the aggressive impulses against themselves: it sublimates them into a cruel blaming of the self, whenever it lets any of these impulses act outwardly. At the same time this idea of agency's freedom conceals its own main function, to conform the individual to social rules and values. The agent prides himself on his autonomy, but is all the while choosing according to the social script.

So agency is a capacity, involving a certain idea of itself. This idea—of an I or self—is on the one hand self-referential, and picks out the capacity. But it also thinks of that capacity not as a capacity but as a (metaphysical) subject. In this situation there are, I think, two options: a) we can identify the I solely by its reflexive role, in which case the term refers to that capacity, or b) we can identify the I by the content the capacity attributes to itself, in which case there is no I.

Nietzsche generally takes the second line, which he expresses in outright denials that any I or self or will ever occurs. 'What separates me most deeply from metaphysicians is: I don't concede that the "I" is what thinks: rather I take the *I itself as a construction of thinking* . . .: so only as a *regulative fiction* with the help of which a kind of constancy and thus "knowability" is inserted into, *invented into*, a world of becoming' (KSA 11: 526 (WLN 20–1)).[32]

And yet it would be just as legitimate to identify the self or I through the self-referential role, hence with the capacity that calls itself so.[33] There really is *something* there, something important. Nor is its idea of itself completely false. Although it lacks the metaphysical status, and the primacy within the organism, that it claims for itself, it is still a significant power within the organism, the most distinctive power humans have. This capacity really is able, in some cases, to control the organism and its (other) drives. And it really does involve a special kind of power, and freedom.

I think that Nietzsche affirms the power and freedom in agency in his account of the sovereign individual in GM II 2. Here are some phrases that seem to apply to this agential control of the drives: 'a true consciousness of power and

[31] EH IV 8: 'The concept of "sin" invented along with the associated instrument of torture, the concept of "free will", in order to confuse the instincts, in order to make mistrust of the instincts second nature!' Also TI VI 7.

[32] Here Nietzsche takes Dennett's view (e.g. 1992) of the self as fictional.

[33] This is Velleman's view, e.g. in (2005).

freedom'; 'this lord of the *free* will'; 'this mastery over himself also necessarily brings with it mastery over circumstances, over nature and all shorter-willed and more unreliable creatures'; '[t]he "free" human being, the possessor of a long, unbreakable will'; 'this power over oneself and fate'. So Nietzsche himself sees reasons to think that a distinctive new power is here, before and without which we were 'slaves of momentary affect and desire' [GM II 3].[34]

Nevertheless Nietzsche is also highly critical of this new capacity. He is critical, we've just seen, of the metaphysical self-conception embedded in our agency. But worse than this ontological mistake is this agency's failure to see its own design, as by and for interests outside. This makes the claim to constitute the organism's essential 'self' false in another way: agency expresses not my individual interests, but social and generic ones. So all its I-ing is a kind of sham, implanted in the interests of taming and herding me. In this respect my drives are the more 'me' than my conscious thinking and choosing. But agency has been designed as an enemy of these drives: it is pitted against the aggressive drives in particular, i.e. precisely those in which people once most enjoyed the 'affect of command', in which they felt themselves most masterly and free.[35] So this agential freedom has been achieved only by a sacrifice of the original freedom in our drives (GM II 17). And to become an individual, *the* individual it is possible for me to be, I must let these drives give me content. I must align my agency with a unity that is achieved in my drives.

So Nietzsche rejects agency's picture of itself as properly independent of the drives. He preaches to agents—his readers—to give credit to their drives, and to learn to dovetail their agency to them. We must hope that our drives have the power—freedom of the first type—to unify themselves, preferably under a single dominant passion. We need to watch for what distinguishes us in our drives and passions, and learn to subordinate our agency—freedom of the second type—to this. So our agency, and the will in our drives, will coincide.

Nietzsche depicts the masters as individuals in which this coincidence of drives and conscious will is achieved. That is, their values, expressed in the principles they affirm and decide by, dictate the same behaviors as their drives. Their values,

[34] I mentioned above that GM II 2 is puzzling, however, and let me now say why. First, Nietzsche depicts this sovereign individual in fully positive terms that give no hint of the limitations I claim he does think this 'agent' has. Second, he describes this individual as 'free again from the ethic of custom, autonomous and supra-ethical', which sounds close to the more Nietzschean freedom I will turn to in c) below. The passage presents this 'supra-ethical' individual as the result of a long social process now complete, whereas I claim Nietzsche thinks he himself only now shows the way to this true independence. My suggestion for resolving this dilemma: that Nietzsche here depicts a *moral* form of agency, which takes its own allegiance to abstract moral principles as a freedom from 'the ethic of custom', i.e. from doing things *as* 'what one does'. But it is still not free from those principles, which themselves enforce social interests in it.

[35] GM II 16: man's taming has involved 'a declaration of war against the old instincts on which his force, desire, and terribleness had thus far rested'.

and their deliberative will, simply express the package of drives they comprise. Nietzsche depicts this as happening naturally and almost effortlessly in them.

It is a puzzle, on my story, how this can be. Why don't the master's values and agency constrain his drives for social purposes, as we've seen agency was designed to do? Where is the asceticism and hostility to drives I've said characterizes agency for Nietzsche? (There is also this other puzzle about the masters: since surely they come after that long bloody history in which 'promising' was bred into us, why do they not suffer from the sickness and gloom Nietzsche says are its products?)

I think the answer must include at least these points. First, the masters belong to the earlier phase of the story GM II tells, the ethic of custom. Guilt—built into the attitude of agency—has not yet become sin before an ultimate God. So it hasn't yet reached its most self-lacerating form, in which the agent views drives as evils in the body. Second the masters' society has been organized to make their own lives possible, with other members constrained to supporting roles (into which their drives are severely forced). But the masters' own role is for war, for which strong aggressive drives will best suit them. So the social interest makes values that favor these drives in them.

But the development and rise to dominance of slave morality—which Nietzsche often just calls 'morality'—has made such masterly harmony between drives and agency much harder and perhaps impossible. Where this morality is strong, it makes agency (conscious thinking and choosing) an avowed enemy to the drives; agency has the task of repressing and indeed eliminating these drives. Of course it fails to notice many of them, and is often used by them unbeknownst to itself—and Nietzsche is constantly pointing out ways these are so. Yet moral agency really does constrain and oppose many of the drives effectively enough to injure them—and in particular the strong drives that would unify action on their own. Nietzsche thinks that we moderns have been drive-damaged by morality's long rule. We have a far richer variety of drives and interests, but these are all—in most of us—weaker than the fewer simpler drives were in earlier ages.

However morality, author of this damage to the drives, is also now under attack, and from one of its own offshoots. Morality, which has shaped the current style of our agency, is very gradually dying the same death that God has. The will to truth that issued out of this morality turned back first against the faith in God, exposing it to the cold eye of scientific thought. This same scrutiny will eventually undermine faith in morality as well.[36] It will pull away such metaphysical supports as those of a soul, surviving death into another world, in which it will find a due reward. And it will gradually make a mockery of the saintly ideal itself—the ideal of a selfless life spent mortifying the drives. This ascetic ideal will (very gradually) lose its power to spur and shape our agency,

[36] See HAH I 34 for an early statement how discovering how our values are rooted in untruth results in one's 'finding his present motives, such as that of honor, absurd, and setting mockery and contempt against the passions which reach out to the future and to a happiness in it'.

how we consciously aim and choose. People feel less and less reason to use their agency against their drives, or indeed to any end at all.

Morality's decline will weaken agency's grip on the drives, and its hostility to them. The way will lie easier, in this respect, to reachieve the healthy fit of agency and drives, which characterized the masters. (We see this in the expanding positive interest in the body and its health and pleasures, which seems an important tendency of the current age.) However there is a different problem. Agency and the drives may compete on a more even footing, but that is by both having sunk in strength. Our drives have been democratized: we have many more of them, a great richness of tastes and wants pulling us in different directions. But rarely is any of them strong enough to be a dominating passion that organizes the rest under it. And agency is also weaker: we feel no moral need to give ourselves a single rule.

Our first task, then, is to re-achieve that synthesis of the two kinds of freedom, freedom in the drives, freedom in agency. 'The *freest act* is that in which our own strongest most finely practiced nature springs forth, and in such a way that *at the same time our intellect* shows its *directing* hand' (KSA 10: 258). For this synthesis, the unifying impetus must come out of the drives, and as agents we must wait and watch for a dominating drive to gel. 'The whole surface of consciousness—consciousness *is* a surface—has to be kept clean from all of the great imperatives. . . . In the mean time, the organizing, governing "idea" grows and grows in the depths—it starts commanding' (EH II 9). Once this drive emerges, conscious agency should give itself over to it.

Goethe is Nietzsche's exemplar of this unity, this double freedom. 'What he willed was *totality*; he fought against the separation of reason, sensibility, feeling, will. . . . Goethe conceived of a strong, highly educated, self-respecting person, skilled in all things physical and able to keep himself in check, who could dare to allow himself the entire expanse and wealth of naturalness, who is strong enough for this freedom' (TI IX 49). (Also WP 933: 'the "great man" is great through the freedom of play of his desires and through the yet greater power that knows how to press these magnificent monsters into service.')

We have, by now, a pretty complicated picture of freedom before us. Yet I think it's still incomplete: there's a third kind of freedom—or element in a comprehensive freedom—which Nietzsche thinks is his main innovation. Goethe achieved that double freedom, that confluence of drives and agency, better than anyone, and Nietzsche highly admired Goethe. But he also aspires, I think, to a kind of freedom that even Goethe didn't have.

c) Self-genealogy. We have so far the idea of freedom as a joint fulfillment of the two kinds of unification we humans have evolved capacities for. The first freedom is drive-synthesis, in particular by the dominance of a single passion. The second freedom is what we call agency, in which conscious thinking and choosing set our course. Both, Nietzsche thinks, are weakening now in our

modern age, so that despite our sense of ourselves as enjoying unparalleled freedom, we have less of either of the two capacities that freedom has so far been.

It's at this point that Nietzsche has his idea of something further and new, his own prospective idea of what freedom can become. He thinks of this as his own contribution and achievement—as the distinctively Nietzschean freedom.[37] It lies in a new kind of agency, an agency redesigning itself for a new purpose. This new agency is, in fact, a continuation and development of the genealogical inquiry we've been pursuing all along. This genealogy of our practices/ideas of freedom not only points the way to the synthesis of drive- and agency-freedom we've seen, it is itself the beginning of a new, third kind of freedom. Genealogy, by allowing agency to at last understand itself—and to redesign itself accordingly—gives us a new way to become a self, a self that feels a new power, and assumes a new responsibility.

Our agency has been trained to hostility against the drives: it fights them as forces foreign and alien to itself—from the 'self' it claims to be. But agency has failed to recognize the way it itself is an 'agent for' foreign forces; it executes a control of the organism by those forces. Its capacities for conscious thought and choice (along with its idea of itself) have been designed by long cultural processes to bind us to morality and social norms, the better to integrate and deploy us. Moreover it expresses the interests of various human types, in various ages, who have redesigned the practice of agency to give themselves what they need (this includes many different kinds of masters and slaves).

Foreign interests are built into our moral principles, built into even the structure of agency, in particular its linguistic structuring (its rules are stated in averaging words). The secret rule of those social interests has opposed what is in fact the basic and indispensable part of the organism, its drives. It's only genealogy that brings these alien interests to light, and gives the agent a chance to oppose and correct for them, as they work in itself. Thus genealogy gives us a new way to be 'free from' control by external wills. It is a way for the organism to take a fuller control over itself, as well as a kind of control over those foreign interests it culls out of itself.

We are familiar with certain ways of being (or not being) free from control by outside forces: political freedom, economic freedom, religious freedom. We have largely overlooked, Nietzsche thinks, a deeper kind of control we are subject to, through our values. GS 335: 'Your judgment "that is right" has a prehistory in your drives, inclinations, aversions, experiences, and what you have not experienced; you have to ask "*how* did it emerge there?" and then also, "*what* is really driving me to listen to it?" ' We need to uncover, that is, the interests that lie behind our values, the functions they have been designed to play in us.

[37] EH IV 7: 'What distinguishes me, what sets me apart from all the rest of humanity is the fact that I *uncovered* Christian morality.'

Nietzsche especially stresses the release from traditional values (though not necessarily genealogy as the route to it) in his conception of the 'free spirit'. Already in HAH I 225: 'what characterizes the free spirit is . . . that he has liberated himself from tradition'; 'he demands reasons, the rest demand faith'. GS 297: 'the acquired good conscience accompanying hostility towards what is familiar, traditional, hallowed . . . constitutes what is really great, new, and amazing in our culture; it is the step of all steps of the freed spirit'. WP 789: 'What a feeling of freedom it is to experience, as we freed spirits experience, that we are *not* harnessed up to a system of "ends"!'

We can't begin to counter the control of these forces and interests unless we can notice them—unless we can see what our values really have been designed to do with us. It is genealogy that exposes these controlling forces. So it is indispensable for the kind of revaluation of these values that Nietzsche has in mind: a revaluation in awareness of the wills embedded in these values. GM P 6: 'we need *a critique* of moral values, *for once the value of these values is itself to be put in question*—and for this we need a knowledge of the conditions and circumstances out of which they have grown, under which they have developed and shifted'.

This insight into the sources of our values begins to make us independent of them. GS 335: 'The insight into *how in general moral judgments ever arose* would spoil these emotional words ["duty" and "conscience"] for you'. It begins to open the affective distance from these values, needed also for genuine revaluation of them.[38]

But how—it must be asked—are we to revalue values set so deeply in us? How can we escape them to find any independent ground from which to assess them? Nietzsche thinks we do have access to a standpoint separate from our agency, and more reliable than it; of course this is the standpoint of our drives. I am to judge my conscious values—and my agential way of living by these values—by how they help or harm the system of drives and passions that I find as my still deeper self. I work to discover the aspects of my agency that are most damaging to my instincts—for example its reliance on guilt—and try to redesign them out of myself. So I learn to remake myself in a way never possible before, a self with a new kind of responsibility for itself.[39]

The point is not, however, simply to stir up and set loose my drives and appetites. Nietzsche is often critical of what he calls '*laisser aller*', letting go.[40] So in TI IX 41: '"Freedom I do *not* mean—"—In times like these, giving over

[38] GS 380: '"Thoughts about moral prejudices", if they are not to be prejudices about prejudices, presuppose a position *outside* morality, some point beyond good and evil to which one has to rise, climb, or fly. . . . One has to be *very light* to drive one's will to knowledge into such a distance and, as it were, beyond one's time; to create for oneself eyes to survey millennia and, moreover, clear skies in these eyes.'

[39] I give sketches of the new 'ethics' and 'politics' this genealogy-based redesign might result in, in (2004), ch. 3 sect. 5–6.

[40] Thanks to Aaron Ridley for reminding me of this further aspect of freedom.

to one's instincts is just one more disaster. The instincts contradict, disturb, and destroy one another; I even define *modernity* as physiological self-contradiction. A rational education would have *paralyzed* at least one of these instinct systems with iron pressure so that another could gain force, grow strong, become master.'[41] Agency will still constrain drives, but now in their own interest: its roles are to discipline them, and to foster the emergence of a unifying passion.

There is another kind of limit to this new kind of freedom. It is also more aware than the earlier kinds were of what it does not and cannot overcome. So it has a more adequate idea of itself. It sees itself as enveloped within necessity and fate, and makes no claim to be a first and uncaused cause. The Nietzscheanly free individual recognizes the contingency in his constitution of drives—there are wills and values here that make him up, and that he can't 'rise above' or 'come before'. And this free one also recognizes his dependence on his moment in the historical drama.[42]

I have spoken of this diagnosis and revaluation of values and agency as something each 'free spirit' is to do for himself. But Nietzsche has in mind that these personal efforts will be cumulative, will together build a new practice and idea of freedom. So he sometimes speaks of this revaluation as carried out at the social- or even species-level. EH III D 2: 'My task, preparing for humanity's moment of highest self-reflection, a *great noon* when it will look back and look out, when it will escape from the mastery by chance and priests and for the first time pose the questions of why? and what for? *as a whole*'.[43]

So genealogy matters not just because it shows us a way to synthesize drive- and agency-freedom, but because it accomplishes this through a kind of agency that achieves a fuller power, selfhood, and responsibility. As I've said this is what distinguishes Nietzsche—or those he makes possible, his overmen—from Goethe, who did not make himself out of insight into his agency and drives. Goethe 'revalues values' by intuitively accepting from the Christian morality only those rules that would favor his drives. Nietzsche revalues values by turning his agency—his conscious and deliberating self—back upon itself so as to understand these values, and the interests they secretly express. By revaluing out of this understanding he controls achievement of that synthesis. He has a freedom—and, we might say, a kind of self—of a sort without precedent even in Goethe.[44]

I take Nietzsche to have this new freedom in mind in his great peroration in GM II 24: 'this bell-stroke of noon and the great decision, that makes the

[41] See also BGE 188 against *laisser aller*: 'slavery is, as it seems, both in the cruder and in the more subtle sense the indispensable means of spiritual discipline and cultivation.' Also WP 122.

[42] Cf. May (1999: 177) on freedom as 'self-responsible commitment to the "necessity" embodied or expressed in one's will[ing]'; May glosses necessity as 'of nature, and specifically of a certain past and "type" '.

[43] EH IV 1: '*Revaluation of all values*: that is my formula for an act of humanity's highest self-reflection, an act that has become flesh and genius in me.'

[44] EH III Z 6 says that 'a Goethe, a Shakespeare, would not know how to breathe for a second in this incredible passion and height' of Zarathustra.

will free again, that gives back to the earth its goal and to man his hope; this antichrist and antinihilist; this conqueror of God and of the nothing—*he must one day come . . .*' By undertaking to step back and judge morality itself, this new individual takes on a new level of responsibility (see e.g. BGE 60 and 212).

This then, I suggest, is Nietzsche's full conception of freedom. Most broadly, freedom is something historical: an ability—with a linked idea of itself—that has been built very gradually through human history, and in such a way that earlier stages are layered beneath more recent ones. To say what freedom 'is' we must tell this history, and also show how this history is now embodied in us, in a layered capacity that works in our drives, in our agency, and now also in our genealogical insight into that agency. Nietzsche argues that with this last step freedom really accomplishes much of what it had claimed to do in agential freedom—it makes one sovereign, and an individual, in senses not yet really true so long as our agency failed to diagnose how its own taming and moralizing design made it work in the interests of foreign forces. Only genealogy lets us understand the design of the values and powers we have taken for granted, opening the way to the new and more adequate freedom Nietzsche commends to us.[45]

REFERENCES

WORKS CITED

Dennett, Daniel (1992). 'The Self as a Center of Narrative Gravity', in F. Kessel, P. Cole, and D. Johnson (eds.), *Self and Consciousness: Multiple Perspectives* (Hillsdale, N.J.: Erlbaum).

____ (2003). *Freedom Evolves* (New York: Penguin).

Gemes, Ken (2009). 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual', (this volume).

Janaway, Christopher (2007). *Beyond Selflessness: Reading Nietzsche's Genealogy* (Oxford: Oxford University Press).

Katsafanas, Paul (unpublished). 'Practical Reason and the Structure of Reflective Agency'. Harvard PhD dissertation.

May, Simon (1999). *Nietzsche's Ethics and his War on 'Morality'* (Oxford: Oxford University Press).

Richardson, John (1996). *Nietzsche's System* (New York: Oxford University Press).

____ (2004). *Nietzsche's New Darwinism* (New York: Oxford University Press).

Velleman, J. David (2005). 'The Self as Narrator', in J. Anderson and J. Christman (eds.), *Autonomy and the Challenges to Liberalism: New Essays* (Cambridge: Cambridge University Press).

[45] I presented earlier versions of this paper at conferences organized by Keith Ansell-Pearson at the University of Warwick, and by Christopher Janaway at the University of Southampton. I am grateful for the invitations, and for the comments by the audiences at these sessions.

This page intentionally left blank

8

Nietzschean Freedom

Peter Poellner

1. INTRODUCTION

The majority of Nietzsche's readers and commentators agree that he rejects certain conceptions of free will or agent causation which claim *either* that, at least sometimes, we are the complete causes of certain events we cause to happen, such that these events happen without causal input from any spatiotemporal entities outside the boundaries of our empirical selves (BGE 21), *or* alternatively that everything among the antecedent and surrounding conditions of an agent's act of will (a volition, a decision to act) could have been exactly as it was, and she might still have decided to act differently (GM I 13). Irrespective of whether Nietzsche's own reasons for being dismissive of 'free will' in either of these senses are sound, many contemporary writers on the subject would concur with his conclusion. If we accept it, as I think we should, it follows that social practices which presuppose such conceptions of free will are rationally indefensible. Nietzsche himself believes that these practices include all forms of morality and of retributive justice which involve the idea that individuals *deserve* favourable or unfavourable treatment for the acts they commit (HA 105; GM I 13).[1]

Yet, there are many places in the later writings where he writes in praise of the 'free spirit' (GS 347; BGE 188), of the 'autonomous' or 'sovereign' individual

[1] For a sensitive discussion and defence of this strand of thought in Nietzsche, see Leiter (2002: ch. 3). It is an interesting question whether the practice of *praising* and *blaming* individuals for their actions or omissions also falls under the purview of Nietzsche's criticism. His own ambivalence—his evident approval of reactive attitudes like admiration and contempt towards individuals on the one hand (e.g. GM I 10–11), and his scorn for the characteristic desire of the slaves and the weak to blame *someone* (whether their masters or themselves) on the other hand (GM III 14–15)—suggests that the issue is not straightforward. This is due to the ambiguity of the terms 'praise' and 'blame'. If what we mean by them is simply an evaluative judgement—perhaps affectively charged—about an individual in account of their actions and dispositions, then there is nothing in principle objectionable about praising and blaming by Nietzsche's lights. If, however, the attitudes referred to by these terms are based on the thought that, among all the causal factors contributing to an

(GM II 2), and of the 'creator of values' who has, in some sense, a 'free will' (GS 347; BGE 213; GM II 2; TI V 3; TI IX 38). Nietzsche has Zarathustra say, for example: 'Creation—that is the great redemption from suffering, and life's easement. But that the creator may exist, that itself requires much suffering and much transformation. [. . .] my *willing* always comes to me as my liberator and bringer of joy. Willing liberates: that is the true doctrine of will and freedom' (Z, 'Of the Blissful Islands').[2]

In order to get a proper perspective on what this 'true doctrine of will and freedom' might be it is important first to clarify what role it is designed to play in Nietzsche's philosophy. His formulations in fact suggest two quite distinct such roles. First, there is the thought that freedom, properly understood, is *constitutive* of being a fully individuated self, a person in the proper sense. This is suggested by his remark that 'we, however, want to become those we are—human beings who are new, unique, incomparable, who give themselves laws, who create themselves' (GS 335). Being a 'unique' human being, rather than merely an instance of a certain type of conscious organism, is here explicitly linked with 'self-creation', and this in turn with giving laws to oneself that one holds oneself to. Elsewhere he speaks of individuals capable of such self-legislation, having an 'independent, protracted will and the *right to make promises*', and possessing their own 'measure of value', as 'autonomous' 'master[s] of a *free* will' (GM II 2). It is plausible to think of Nietzsche as concerned here with what might be called a *transcendental* question: the constitutive conditions of full-fledged, autonomous rather than heteronomous, selfhood.[3]

In other passages, what 'freedom' and being a 'free spirit' gesture towards seems more appropriately described, not as a condition of being a self in the proper sense, but as a *substantive ideal*. This is arguably what is at stake in many of those remarks where Nietzsche expresses admiration for people who, as he sees them, have succeeded in integrating an unusually great multiplicity of 'drives' and evaluative commitments into a long-lasting, coherent whole: 'a philosopher . . . would be compelled to see the greatness of man, the concept "greatness", precisely in his spaciousness and multiplicity, in his wholeness in diversity: he would even determine value and rank according to how much and

action, only those lying within the individual agent are evaluatively relevant, *and* that the agent therefore merits, in addition to the judgement, favourable or unfavourable treatment (reward or punishment) by virtue of her action, then Nietzsche's criticism applies. Praise and, especially, blame in this second sense are central constituents of the 'slave morality' excoriated by him (GM I 13; TI VI 7).

[2] Nietzsche's works are referred to by their standard English abbreviations. Translations are based on the English editions listed in the bibliography, sometimes with emendations. *Nachlass* material not included in *The Will to Power* is cited from Nietzsche's *Werke: Kritische Gesamtausgabe* (KGW), ed. G. Colli and M. Montinari, who have also edited Nietzsche's letters as *Briefwechsel: Kritische Gesamtausgabe* (KGB). Translations from these editions are my own.

[3] See Guay (2002) and Gemes (this volume).

how many things one could endure and take upon oneself. . . . He shall be the greatest who can be . . . the master of his virtues, the superabundant of will; this shall be called greatness: the ability to be as manifold as whole, as vast as full' (BGE 212; cf. GS 290; TI IX 49). This conception of greatness as a certain kind of 'mastery' and 'superabundant will' is one of the most recurrent themes in the later Nietzsche's multifarious attempts to determine an order of 'value and rank'. And the ideal of 'unity in maximal diversity' is complemented and further specified by others which also employ the vocabulary of freedom, creation, and mastery. Let me mention, without claims to exhaustiveness, a few of these.[4] Sometimes Nietzsche is drawn to the view that the most valuable kind of 'freedom' would be one that is expressed by way of a constant and successful struggle and 'effort' against strong internal resistances: 'One would have to seek the highest type of free man where the *greatest resistance* is constantly being overcome' (TI IX 38, my emphasis; cf. Z, 'Of Self-Overcoming'). Elsewhere, he stresses the free spirit's lack of need for certainties (GS 347), or for explicit rules and imperatives, insisting 'that every action that has ever been done was done in an altogether unique and irretrievable way, and this will be equally true of every future action; that all regulations about actions relate only to their coarse exterior (even the most inward and subtle regulations of all moralities so far); that these regulations may lead to some semblance of sameness, *but really only to some semblance*' (GS 335). According to these passages, the free spirit is an ethical particularist, realizing that the features potentially relevant for evaluative appraisal of an action are so manifold, fine-grained, and mutually modifying that no substantive generalizations, capturing all these features and yet also being equally applicable across different actual contexts, are to be had.[5]

[4] What follows is only an incomplete list of some core aspects of the later Nietzsche's ideal(s) of human value. I am leaving aside here, for example, his advocacy of the 'virtue of bestowing' (e.g. Z, 'Of the Bestowing Virtue'), of *amor fati* (e.g. EH, 'Why I am so Clever', 10), and—one of his most difficult thoughts—of a projected overcoming of the human: 'I regard all metaphysical and religious modes of thinking as consequences of a dissatisfaction *with man*, of a drive towards a higher, supra-human future—except that people wanted an escape to a beyond for *themselves*: instead of building the future' (KGW VII.2. 27.[74]; also Z, 'Prologue', 4). 'Nothing can be predicted, but through a certain elevation of the type *man* a *new* power could reveal itself of which we have known nothing so far (namely a synthesis of opposites?)' (KGW VII.2.34.[125]). These latter ideas, which occur in various modulations in *Thus spoke Zarathustra* and in many notebook entries of the period 1884–5, *prima facie* stand in strong tension with his praise elsewhere of an intermittently already *attained,* Goethean-humanist 'naturalness' (TI IX 49) and satisfaction with oneself (GS 290). Another central aspect of Nietzsche's substantive ideal will be discussed in Sect. 3 of this essay.

[5] This interpretation of Nietzsche's reflections on the uniqueness of individual actions and on the inadequacy or uselessness of 'even the most inward and subtle regulations' is indebted to Robert Musil's development of Nietzsche's point in *The Man without Qualities*: 'The value of an action or of a quality, indeed their essence and nature, seemed to him dependent on the circumstances surrounding them, on the ends that they served, in short, on the whole complex—constituted now

The criteria of adequacy for Nietzsche's construal of freedom will clearly be different depending on whether what is at stake is an evaluative ideal or the conditions of the possibility of fully individualized—'unique'—selfhood (although the satisfaction of these very conditions may of course also be *part* of or presupposed by a normative ideal). One of my claims in this essay will be that *formal* features of comportments, preference systems, or motivations—important though they are—are not by themselves sufficient on either score. For example, if we take Nietzsche to be primarily concerned with an evaluative ideal, we might be tempted to combine some of the formal characteristics I have listed above—say, a long-lasting, 'protracted' unity among a maximal diversity of 'drives', freedom from the need for certainties and from explicit moral rules—and identify Nietzsche's ideal with this combination of features. Yet it seems clear that a person might exemplify all these characteristics and yet appeal neither to us—even the most incontestably free-spirited among us—nor to Nietzsche. Such a person might have shaped, perhaps by way of continued reflection, self-criticism, and against various inner resistances, a remarkable unity out of a great diversity of his actual drives or desires, individuated by their contents—desires which might include, as it happens, the desire to watch Reality TV shows, the desire to download jingles from the internet, the desire to avoid conversations other than small talk, and so forth in a similar vein—and he may act in each situation in a thoroughly particularist manner without relying on pre-formed rules. One only needs to give brief attention to this frivolous instance of unity-in-diversity to concur that, when interpreting Nietzsche's remarks on these issues, we implicitly credit him with being concerned mostly with significant, deep, or, in his language, 'higher', desires and motives. But once this is conceded, it becomes evident that, necessary as the formal characteristics delineated so far may be for Nietzschean freedom, an independent specification of what makes a motive significant or 'higher' is required to supplement them. As Alexander Nehamas has remarked, with similar points in mind: 'The existence of character may not be quite as independent of the *quality* of the actions of which it constitutes the pattern'.[6] In what follows I shall argue that Nietzsche considers two kinds of non-formal, 'qualitative' or substantive commitments to be necessary—albeit not by themselves or even jointly sufficient—for freedom understood, respectively,

thus, now otherwise—to which they belonged. . . . Then all moral events took place in a field of energy the constellation of which charged them with meaning, and they contained good and evil just as an atom contains the potentialities of chemical combination. They were, so to speak, what they became, and just as the one word "hard" describes four quite different essences according to whether the hardness relates to love, brutality, eagerness or severity, so the significance of all moral happenings appeared to him the dependent function of others. In this manner an endless system of relationships arose in which there was no longer any such thing as independent significance, such as in ordinary life, at a first crude approach, is ascribed to actions and qualities' (Musil: 1995, i. 297–8).

[6] Nehamas (2001: 276–7). It is one virtue of Leiter's account of Nietzsche's commitments that he is quite explicit about some of their substantive, non-formal features. See Leiter (2002: 115–25).

as a condition of full personhood and as a normative ideal. I shall first, in Section 2, turn to the 'transcendental' question of freedom and personhood, and discuss one central component of his evaluative ideal of being a 'free spirit' in Section 3.

2. FREEDOM AND PERSONHOOD

For Nietzsche, the experience—and not merely the thought[7]—of effective *agency* plays a constitutive role for potentially self-conscious subjectivity:

> The degree to which we feel life and power . . . gives us our measure of 'being', 'reality', not-appearance. The subject: this is the term for our belief in a unity underlying all the different impulses of the highest feeling of reality. (WP 485; cf. WP 488).

This experience of self-efficacy requires 'resistances' or 'obstructions' (WP 689, 693, 702) and the notion of objective reality is derived, in part, from this experience of resistance to volitional agency. The concept of agency is linked to that of desire, while desire, in the sense relevant here, as a personal-level wanting to maintain or to change some experiential content, depends on what Nietzsche calls 'affects' (more on which below). The upshot of this Nietzschean line of thought is condensed in the famous passage on perspectivism in GM III 12, where he claims explicitly that the concept of objective reality is dependent on affectivity: 'the *more* affects we allow to have their say on a matter . . . the more comprehensive will be our "concept" of the thing, our "objectivity"'. So, according to Nietzsche, subjectivity, in the sense of a subject's having a conceptual grasp of the distinction between *self* and *world*, requires the experience of agency, and the latter presupposes affectivity. But the concept of agency here is rather thin. It is plausible to ascribe experienced agency, and the corresponding level of subjectivity, even to very young children or to adults suffering from compulsive syndromes or strong addictions—individuals who, in Nietzsche's terms, are 'slaves of momentary affect and desire' (GM II 3).

[7] Nietzsche's point here invites comparison with Kant's conception, in the *Critique of Pure Reason*, of transcendental subjectivity as consisting, in each case, in a subject's consciousness of herself as a 'self-active' being synthesizing representations. *Qua* transcendental subject 'I exist as an intelligence which is conscious solely of its power of combination' (B 158). The nature of this 'consciousness' is notoriously unclear in Kant. He says that 'I do not have another self-intuition which gives the determining in me (I am conscious only of the spontaneity of it)'. Hence 'I cannot determine my existence as that of a self-active being; all that I can do is to represent to myself the spontaneity of my thought'. Kant's formulation here is ambiguous: is the 'representation' of my synthesizing spontaneity only a *thought*—essentially an *inference* from the unified contents of consciousness, as B 134–5 suggests—or is it experientially ('intuitively') presented? In which case Kant's denial of self-intuition as an 'active being' would refer not to the spontaneity itself, but only to the *intrinsic* character of that which exercises the spontaneity, revealing Kant's commitment to a metaphysics of substances with intrinsic properties at the noumenal level (see esp. B 478).

Mostly, Nietzsche seems to be concerned with agency and subjectivity in a more substantial sense. He is interested, like Kant, not just in minimal agency, but in autonomous or free agency in a sense which it would be inappropriate to ascribe to a compulsive subject, or to an addict or acratic person—a slave of momentary affect and desire—or to a very young child. Some of the best-known places where Nietzsche extols autonomy or free agency are his reflections on the 'sovereign individual' in GM II 2, and on self-knowledge, self-creation, and the free spirit in GS 335 and GS 347. What emerges from these passages is that autonomy or freedom in the relevant sense is a matter of 'having a protracted will' and 'mastery over oneself' under the aegis of a 'conscience' (GM II 2). The autonomous individual experiences her effective desires and her actions as originating from and *expressing her self*: 'We, however, want to become those we are—human beings who are new, unique, incomparable, who give themselves laws, who create themselves' (GS 335). It seems clear that such a capacity for self-creation requires a minimal reflective capacity, an ability to detach oneself from one's impulses, so scrutinize, modify and revise them in the light of rational considerations. Without these abilities, there could be no *self-imposed* laws as envisaged in this passage. For Nietzsche understands the self as creating itself, as coming into being, *through* giving rules or laws to itself which it holds itself to by a 'protracted will' (GM II 2). We should not think of these 'laws' as explicit, symbolically represented rules, and Nietzsche's remark about the uniqueness of every action and the coarseness of every explicit rule suggests that he does not think about them in this way. But what is essential to someone's giving a law *to herself* is her ability to *recognize* what constitutes acting in accordance with it and what constitutes violating it. Hence, she must be able to appreciate certain states of affairs as *reasons* for behaving in one way or another. Indeed, giving oneself a law is virtually synonymous with acting in the light of, and not merely in accordance with, reasons. An organism that is not capable of appreciating reasons from a first-personal perspective may exemplify behaviour that is lawful (as well as being necessarily unified in various ways), but the laws it obeys cannot be *self*-given; and since self-creation requires this self-imposition, such organisms cannot be *selves* at all.[8] The appreciation of something as a reason for acting one way rather than another implies at least the minimal reflective capacities alluded to above: the ability to recognize reasons-pro and

[8] Similar points are made by Guay (2002). What is problematic in Guay's interpretation is his further claim that Nietzsche can say nothing substantive about what might constitute good and bad reasons; good reasons simply are those that we continue to find compelling in the light of ongoing reflection and self-assessment. But if anything is characteristic of Nietzsche's thought about these matters—whether early or late—it is that the difference between what actually are good reasons and what insistently appear to be good reasons to the agents themselves is not abrogated, in principle, by moving to the collective level or to the grand historical perspective. Dishonesty with oneself, *ressentiment,* and the rest of Nietzsche's *bêtes noires* might run very deep in a collective form of life and might conceivably turn out to be ineradicable by the methods of inquiry that happen to be available within that form of life.

reasons-con, and to distinguish, with awareness, one possible course of action from another. It therefore implies self-consciousness.[9] The claim that laws can only be self-given if an agent has an ability to reflect on reasons guiding her actions should be distinguished from two other, superficially similar theses, which Nietzsche is not committed to. It is clearly not Nietzsche's view that the free agent *acquires* most of her self-constitutive motives, desires or beliefs in a reflective or deliberative way. On the contrary, he tends to emphasize the role of social constraints, 'custom', and hereditary factors in the aeteology of freedom (e.g. GM II 2; BGE 213). Nor is Nietzsche claiming that the free agent is especially prone to reflecting on her commitments or questioning them. What is essential is only this: that the agent be *capable* of consciously distinguishing between actual or possible motives, courses of action, etc., she takes as expressive of her self and others that she does, or would, not take to be so.[10]

But what makes a motive suited to be expressive of who I am, to be an instance of 'self-creation' in Nietzsche's sense? His polemics against the Kantian concept of autonomy in GS 335 and elsewhere make it clear that the relevant feature cannot be the universalizability of the motive's content in a Kantian sense. In the same passage, Nietzsche gives us a clue to an answer when he says that 'our opinions, valuations, and tables of the good certainly belong among the most powerful levers in the mechanism of our actions. . . . Let us therefore limit ourselves to the purification of our opinions and *valuations* and to the creation of our own new *tables of what is good*' (GS 335; my emphases). Nietzsche's accentuation here of a subject's commitment to 'tables of the good', to evaluations and their purification recalls contemporary compatibilist accounts of freedom which tend to contrast mere desire with evaluation.[11] But what is it to *value*, as opposed to *merely desire* something? The contrast here clearly cannot be merely one between first-order desires and higher-order desires to be motivated by certain first-order desires, for such hierarchical accounts cannot by themselves explain what we need an explanation of, namely what makes a particular second- (or higher-) order desire *mine*.[12] Now, Nietzsche does have a theory of evaluation and its grounds, but it is a theory which at first sight seems to undercut the very distinction between evaluation and *mere* desire. For according to him, our basic mode of acquaintance with value is through 'affects' (*Affekte*). Thinking of things in evaluative terms

[9] See Sect. 3 for an interpretation of Nietzsche's view concerning the relation between a consciousness of reasons and the subpersonal physical processes subtending it.

[10] I am grateful to Ken Gemes for objections forcing me to clarify these distinctions.

[11] See e.g. Watson (1982) and Taylor (1989: ch. 1). Taylor contrasts desire with 'strong evaluation', expressed in judgements of, and qualitative distinctions between, desires as 'higher' or 'lower': 'a good test for whether an evaluation is "strong" in my sense is whether it can be the basis for attitudes of admiration and contempt' (1989: 523). Nietzsche's evaluative practice is eminently one of strong evaluation in this sense. Not only is the higher/lower distinction omnipresent in it, so are admiration and contempt (see below).

[12] This is a point rightly stressed by Leiter (2002: 94). Cf. Watson (1982: esp. 107–9).

rationally requires thinking of them *as* having a bearing on an actual or possible affectivity:

> Every ideal presupposes *love* and *hatred, admiration* and *contempt*. Either the positive emotion is the *primum mobile* or the negative emotion. For example, in all *ressentiment* ideals *hatred* and *contempt* are the *primum mobile*. (KGW VIII.2.10.[9])
>
> . . . moralities too are only a *sign-language of the affects*. (BGE 187)

How are we to interpret Nietzsche's talk of affects as conditions of the possibility of moral and other evaluative 'ideals' in passages like these? Examples of affects in his sense are hatred, envy and greed (BGE 23), as well as courage, shame, and anger (KGW VII.1.7[87]). All of these states would normally be classified as *emotions*, and Nietzsche's examples indicate that this is the central meaning of 'affect' operative in his account, although the inclusion of a 'desire to dominate' (BGE 23) among his illustrations suggests that he also uses the term more broadly to refer to any kind of mental state with a distinctive phenomenology of favouring or disfavouring, of attraction or 'repulsion' (ibid.).[13] His general thought, then, seems to be that our most fundamental acquaintance with values is through affective states, and in particular through emotions. Such an approach accords well with the familiar observation, traditionally stressed by ethical expressivists, that sincere categorical value judgements necessarily express a *commitment* of the subject making the judgement. However, this view, which has strong independent philosophical credentials, prima facie conflicts with another claim of equally strong standing: that the contents of such judgements, and perhaps also of evaluative experiences themselves, have a kind of *objectivity*. The idea of objectivity alluded to here is very general and minimal; it is simply the notion of a *constraint* upon impulses which makes these intelligible to the subject herself and to others *as preferences*. As has often been noted, the very idea of a choice among possible alternatives of action is tied to that of constraint in just this sense: there must be criteria available to the choosing subject—although

[13] This characterization of affect as having a phenomenology might be thought to fall foul of Nietzsche's countenancing of unconscious mental states, including affective ones (e.g. GS 354, 357). However, we are here concerned only with affects in so far as they as constitutive conditions of value. There are powerful independent arguments for the contention that the concept of value is *essentially* a phenomenological concept—it picks out, in its logically most basic use, instantiations of phenomenal properties (see Section 3). As regards the exegetical question of how best to interpret Nietzsche's remarks on unconscious mental states, it seems clear that Nietzsche is not denying that such states at some level have phenomenal properties (a certain what-it-is-likeness). Indeed, he is often attracted by the view that *everything* actual necessarily has such properties: 'What is "appearance" for me now? Certainly not the opposite of some essence: what could I say about any essence except to name the attributes of its appearance! Certainly not a dead mask that one could place on an unknown x or remove from it!' (GS 54; cf. WP 556–69). The textually most compelling reading of Nietzsche on unconscious mental states is that many episodic mental states are inferentially unavailable to the mind's main-system, displaying no phenomenal properties to *it*, while yet possessing such properties (see also KGW VII.3.37.[4]). For a more detailed defence of this interpretation, see Poellner (1995: ch. 5.2). For the distinction between mental main-systems and sub-systems, see esp. Pears (1985).

these need not be explicit—which make it appear to her that one course of action is better than another. In the absence of such criteria we cannot speak of choice at all.[14] But once the idea of criteria for action is in play, then so is the notion of error. A subject that is capable of making choices in the light of criteria is also, in principle, liable to failure in responding to the presence of these criteria, for example by failing to notice features of a situation that she would herself regard as relevant. As we saw above, Nietzsche himself is very much alive to this normative element of minimal objectivity presupposed by free agency when he speaks of 'laws' which can guide an individual's actions and which therefore also allow for the possibility of actions recognizable by the agent as contravening them.

How are we to understand this element of minimal objectivity or normativity required by free agency? Given that Nietzsche rules out a Kantian understanding of the laws that guide the free person's actions (GS 335; BGE 187; GM II 6), and given that all evaluation is considered by him as grounded in 'affects', it would seem that the affects themselves are seen by him as capable of providing normative constraints. Indeed, it is tempting to read his repeated references to 'taste' as pointing precisely to a structured pattern of conscious affectivity which itself embodies and expresses such normative constraints. Nietzsche says: 'I have a "taste" . . . , but no reasons, no logic, no imperative *for* this taste'[15]—that is, 'taste' constitutes the most basic level of reasons for action.

The idea that affects, 'taste' or 'sentiments' can ground evaluations which allow for a minimal objectivity of the utterances expressing these evaluations, and thus make possible bona fide evaluative *judgements*, is of course not unique to Nietzsche. In recent years, this idea has been associated especially with the metaethical approach Simon Blackburn calls 'quasi-realism'. Blackburn agrees with Kant that what is foundational for the component of objectivity in ethical judgement is the experience of *obligation*.[16] But, *pace* Kant, it is not the awareness of a categorical prescription or imperative that gives rise to this, but those 'sentiments' (affects, in Nietzsche's language), which happen to be most deeply entrenched in the psychology of the subject who feels herself under an obligation. Owing to a particular social conditioning, a person may have strong feelings, interpreted by her, say, as feelings of filial affection and loyalty, which in certain circumstances (e.g. a parent being ill and in need of care) produces a second-order, reflective psychological constraint—a feeling of obligation—counteracting and capable of overriding conflicting first-order desires (e.g. the desire to leave home and strike out on her own). In

[14] For a classical criticism of the incoherence of the notion of 'radical choice' not guided by criteria, see Taylor (1982).

[15] Emphasis mine. Letter to Heinrich Köselitz (Peter Gast), 19th November 1886 (KGB III.2, 284). Cf. GS 39 and Z, 'Of the Sublime Men'. For comments on the relation between conscious taste and *physis* thematized in GS 39, see Sect. 3 below.

[16] See Blackburn (1981: 175–8), and (1985: 8).

this way, sentiments themselves may give rise to normative constraints and guide action. Yet, in Blackburn's quasi-realist picture, those sentiments most deeply lodged in an individual's psychology, which give rise to the sense of obligation in relevant circumstances and provide the normative discipline, are Humean, logically distinct, *effects* of the 'object' (the parent) or of the subject's relation to the object. The sentiments are not *intrinsically about* the object and can thus be described in their phenomenal character without referring to any properties of the object. Thus, on this account, sentiments with type-identical intrinsic phenomenal characters could be experienced in relation to distinct objects, persons, or states of affairs that are not represented by the subject as having any properties in common, other than their causing the same sentiment.

In his middle-period writings Nietzsche occasionally makes remarks which evince a similar psychology of the affects, but on closer reading his position even at that time is more ambivalent on the matter of the 'subjectivity' and putative world-independence of the intrinsic character of conscious affects:

> One loves neither father and mother, nor wife and child, but the pleasant sensations they produce in us. . . .
>
> If the representation of God is removed, then so is the emotion of 'sin' as an offence against divine precepts. . . . Probably what remains then is the sort of displeasure which is fused with, and related to, the fear of being punished by wordly justice and of being disrespected by people. (HA 133)

While the first sentence may suggest something like the Humean picture of the sentiments, the remainder of this passage clearly gestures towards a different construal of emotions as *co-constituted* in their phenomenal, experienced character by *representations* of the world or aspects of it. As we shall see below, this view is more unequivocally developed in later writings.

The Humean, and quasi-realist, account of conscious affects as merely causally, extrinsically, related to the world faces serious problems, which it has never been able to address satisfactorily. Most importantly, it is phenomenologically inadequate as a description of what actually goes on in much ordinary affective-evaluative experience. Symptomatic of this descriptive inadequacy are its difficulties in accounting for the differences between such everyday experiences and those of subjects whom we might want to call self-alienated—who experience their most firmly entrenched affective responses as irrepressible conditioned reflexes or as quasi-compulsive. The problem here is not resolved by insisting that the normal subject *identifies* with her firmest affects, such that her self-image or self-respect is somehow bound up with them, while the self-alienated subject does not or cannot do this, for this is no more than a redescription of the problem. What is at issue is precisely how it is possible for a subject to identify herself with some of her affects if, on reflection, she is rationally constrained to regard them as not potentially disclosing any features of the

world, but as non-representational, sociologically fortuitous conditioned reflexes. It will obviously not do to respond to this that the subject may identify with some of her affects by having learned to *infer* that they are reliable indicators of what she values—for this would require that what she values is disclosed to her independently of them, which is what both Nietzsche and the quasi-realist Humean deny.

No doubt an individual may be entirely untroubled by, and even satisfied with, what in his more reflective, theoretical moments he recognizes to be an adventitious, conditioned, non-representational response-pattern he happens to have had inculcated into him—just as it is perfectly possible to be untroubled by and satisfied with some quirky behavioural trait one happens to have acquired. This kind of self-relation is just what self-conscious, and genuine, 'excentricity' consists in. What is characteristic of such excentricity is not necessarily individual idiosyncrasy—there may be communities of people who are excentric in quite similar ways—but rather a subject's awareness that her comportments or preferences reflect and express fortuitously acquired subjective dispositions not objectively 'appropriate' or answerable to anything in the world. This kind of self-relation—which characteristically tends to involve a heightened, self-objectifying, form of self-consciousness—is clearly different from the relation philosophers have in mind when they speak of a subject's identifying herself with (some of) her feelings, desires, or actions. If such excentricity extends to a subject's central or deepest commitments and orientations it comes to resemble those instances of putative identification with some bizarre set of desires that can be produced by conditioning or 'brainwashing', sometimes illustrated in contemporary discussions by fictional cases like *Brave New World* or *Clockwork Orange* (see Watson, 2004: 170–3). Our intuition is that subjects who have been brainwashed into 'identifying' with some excentric range of desires lack freedom. Gary Watson suggests that what explains this intuition is the fact that such subjects are 'incapable of effectively envisaging or seeing the significance of certain alternatives, or reflecting on themselves and on the origins of their motivations, of comprehending or responding to relevant theoretical or evaluational criteria' (ibid: 172). But we want to say that in such cases of evaluative conditioning, the subjects are unfree even if their observational and inferential capacities in respect of non-evaluative features of the world around them are unimpaired, and if they are quite aware of the causes of their motivations.[17] And to suggest that they do not respond appropriately to relevant *evaluational* criteria simply raises once more the question of what warrant we have—not just as immersed, non-reflective agents, but also in the light of theoretical reflection—for regarding

[17] Note that someone's failure to register instantiations of *hedonic* properties such as pain and pleasure is not a cognitive deficiency with respect to *non-evaluative* features of the world (e.g. to physiological or functional properties). It is a failure to pick up experiential, phenomenal features that are essentially value-involving. On the relevant concept of pain, see Strawson (1994: 247–50), and Scheler (1973: 105).

the brainwashed subject's evaluational responses as deficient, if all evaluations are based on affects, and if all affects are on a par in having causes while not constitutively involving representational reasons.[18] Nor is the difficulty resolved by suggesting that the normal subject's most deeply entrenched attitudes are such that she cannot even *envisage* what an improvement of these attitudes might consist in; for such inability is quite typical of any successful deep conditioning.

It is difficult to avoid the conclusion that non-representational accounts of conscious affects do not have the resources to explain convincingly why we consider the evaluative 'identifications' of brainwashed subjects to be expressive of unfreedom—of a lack of autonomy. In his later writings Nietzsche suggests a different way of understanding the normative component of affectivity. A recurrent idea in these later texts is that what underwrites the talk about the objectivity of at least some values is not, or at least not fundamentally, the second-order mental state experienced as a sense of *obligation*, but rather the nature of some of our first-order affects themselves. What is characteristic of the emotions we are inclined to describe as love, admiration, or contempt, is that they are normally experienced not merely as *caused* by their objects, but as *merited* by them. In undergoing such emotions, a subject experiences an affective attraction or repulsion which seems to be exercised upon her by the object itself in virtue of some property the object has, such that the essential nature of the emotion could not be adequately specified without reference to this (apparent) property of the object. In other words, the affective response is itself experienced as an *appropriate* response to some feature of the object, as a picking up on some value-aspect pertaining to the object. As we shall see, Nietzsche warns against misconstruing this type of affective experience of apparent objective evaluative properties of an object—say, of a painting's beauty as presented *in* the experience of aesthetic admiration, or of meanness of character as presented *in* the experience of contempt—as having any *metaphysical* significance. Affective experience itself does not commit us to any view about whether or not what is presented in it has the right metaphysical credentials. But, Nietzsche suggests, it *does* commit us to ascribing at least *phenomenal objectivity* to those features which it appears to the subject to be a registering of.

[18] If one were to respond to this objection by insisting that the question of evaluative justification cannot arise at the theoretical level, but only for the immersed participant, then this simply runs counter to the actual structure of evaluative practice and discourse. For the latter standardly does not recognize a distinction of this kind, but regards the question of the possible *justification* of even the most deeply entrenched affective responses—beyond just emphatically reaffirming them and pointing to their causes—as arising in both contexts. The picture of the human agent as being systematically divided between a level of immersed agency and practice, and a level of (at least rudimentary) philosophical reflection, such that the results of the latter can have no bearing on the former, is not only unfaithful to most actual (non-excentric) ethical practices, it is of course also un-Nietzschean.

The much-cited analogy of the experience of colours is illuminating in this particular respect, although not in others. When I perceive a lemon's being yellow, I seem to be perceiving a property of an object, and my normal colour experience, and indeed my visual experience of physical objects, could not be described or even rendered intelligible without construing it in this way. To say that, when perceiving the lemon's being yellow I am perceiving a property of an object, does not necessarily carry metaphysical implications; perhaps colours as we perceive them are not features of the world as it is in itself. Yet, my experience here is incontestably of a *phenomenally objective* property. An item is objective, rather than subjective, in this sense, just in case it is not identical with any one experience of it, or with any part or aspect of an experience. Rather, it is presented as being available for many numerically distinct experiences of it by oneself and by others.[19] It is objectivity in this modest sense that is essentially implicated in the ordinary, everyday concept of perception. Now, Nietzsche proposes that many 'affects' and the evaluations they involve should be construed quite literally as *perceptions* of phenomenally objective features of the objects they are directed at:

> Master morality *affirms* just as instinctively as Christian morality *denies*. . . . These contrasting forms of the optics of value are *both* necessary: they are ways of seeing which are unaffected by reasons and refutations. One does not refute Christianity, just as one does not refute a defect of the eyes. (CW, 'Epilogue')

In *The Gay Science* he makes the same point, with some additions:

> [The higher human being] is always haunted by a *delusion*: He fancies that he is a *spectator* and *listener*, confronted by the great visual and acoustic spectacle of life; he calls his own nature *contemplative* and overlooks that he is also the real poet who keeps creating life. . . . We, the thinking-feeling beings [*die Denkend-Empfindenden*], are the ones who really continually *make* something that was not there before: the entire ever growing world of valuations, colours, accents, perspectives, scales, affirmations, and denials. . . . But precisely this knowledge we lack . . . (GS 301)

The complex point emerging from both passages, taken together with his more general claim about the affect-dependence of evaluations, including those systematized in 'moralities', is this. The affects that lie at the basis of evaluations (in the sense in which these contrast with the mere 'momentary . . . desire' of GM II 3) are *representations*; in fact they are, phenomenologically, perceptions in the everyday, metaphysically undemanding sense outlined earlier. And it is precisely the fact that they are experienced as perceptual by their subjects

[19] For this concept of phenomenal objectivity, see e.g. McDowell (1985) and Dancy (1993: ch. 9). An item is *phenomenal* in the sense relevant here if it has some qualitative what-it-is-likeness, i.e. some phenomenal property. Subjective states can have phenomenal properties (e.g. the pain caused by a pin-prick), but so do the objects of our everyday life-world (the yellowness of a lemon).

that *explains* the 'delusion', to which even the 'higher human being' is prone, that the values represented through them are there in the world anyway, quite independently of any such affective states. The latter point is of course just what Nietzsche denies: 'Valuations, colours, accents' are 'created' by our conscious engagement with the world; they do not exist absolutely, entirely independently of this engagement. Yet, while values are thus, in one sense, *created* by us, they are, in another sense, *discovered* in our affective encounter with the phenomenal world, which is the world of our everyday lives.

At this point, it may be helpful to retrace some of the steps of Nietzsche's argument so far. His account of freedom starts with a distinction between self-given 'laws' (GM II 2) and 'valuation' (GS 335) on the one hand, and 'momentary affect and desire' (GM II 3) on the other. This raised the question of how such a distinction can be sustained if valuations themselves, as Nietzsche argues, are based on affects. While it is clearly a strength of this kind of 'emotivist' approach to valuation that it can easily account for the necessarily commitment-involving character of sincere categorical value-judgement, it finds it more difficult to make room for the element of minimal objectivity (normativity) essential to valuation. I mentioned one familiar proposal in this context, according to which the element of normativity might be provided for if we can make sense of an individual's being in a position to understand some, but not all, of her possible affects, as her *own* and, in this sense, to *identify* with and thus potentially to 'appropriate' them. Yet, the traditionally dominant accounts of the affects (emotions) as non-representational effects of external or endogenous stimulation seem to make such identification, in any sense which would be relevant to an individual's *autonomy*, impossible.

The alternative idea that emerges from Nietzsche's later construal of (at least some) emotions as perceptual is this: if we can rightly understand some affects as representational states, we can both make sense of the normativity of evaluation and of the idea of identifying with one's affects, of experiencing them as one's own, rather than as brute impulsions, whether innate or conditioned. Affects experienced as one's own, as expressing what one is, would on this construal be those one understands, often pre-reflectively, as being *appropriate*, as answerable to and thus as actively relating oneself to, the world. Precisely in being representational, affects would be, in Kant's terminology, functions of spontaneity (activity), rather than merely passive effects, and thus valuation would engage, even at the most basic level, self-activity (just as sense-perception does). By contrast, if all valuation-generating conscious affects were themselves merely passive effects, no higher-level reflective ordering or disciplining them would be capable of making *them* express or 'create' one's self—freedom would not go 'all the way down'.

Yet, the idea of (some) emotions as perceptual states is obviously not without problems. The relevant everyday concept of perception—which, to repeat, is metaphysically neutral—includes the following essential features: (i) perceptions

have representational objects, they are *about* something; (ii) while thus having intentional content, they are neither judgements nor dispositions to judge; (iii) they are epistemically direct representations; (iv) they systematically appear caused by the objects/features they represent; (v) they represent phenomenal properties of their objects; (vi) their conditions of success include veridicality. Space does not allow me to examine here whether and how emotions might exhibit all these characteristics. Concerning (i) to (v), I can merely assert here that they do not, in my view, present insuperable obstacles to a perceptual account of some emotions.[20]

Arguably the most problematic aspect of the perceptual account of affectivity suggested by Nietzsche is feature (vi) above—the potential veridicality or truth-aptness of perceptual contents. Can emotions, like sense perceptions, intelligibly be said to be, on occasion, *veridical*? The idea that emotions are normatively assessable as appropriate or inappropriate is of course a standard theme in contemporary philosophy of the emotions.[21] There is also widespread agreement that conscious emotions essentially are or include evaluative appraisals of their objects.[22] But is it permissible to construe the appropriateness of an emotion as the *veridicality* of its evaluative content? This is precisely what seems to be implied by Nietzsche's talk of the affects underlying Christian morality as involving, metaphorically, a 'defect of the eyes' prior to any reflectively considered or inferential 'reasons' (CW, 'Epilogue')—that is, a *cognitive* deficiency at the most basic level of encountering value.

I want to develop this point by considering some of the affects expressed in Nietzsche's own texts as examples. Nietzsche's own evaluative practice typically consists in, or includes prominently, the expression of such affective responses—this is one peculiarity, which makes his texts so different in style and content from more standard philosophical writings. Most frequently, the emotional responses whose expression is recorded in his writings are directed at *other* conscious affective states of himself or of other (actual or fictional) people. A central case in point here is his revulsion at, and contempt for, *ressentiment,* recorded in Essay 1 of the *Genealogy* (see esp. GM I 10–14). Nietzsche makes it clear that affective valuations such as this are not, or should not be, primarily directed at one's own or others' behavioural dispositions or actions, but at the affective structures motivating actions (BGE 268, 287). Human value is indeed largely identified by him with those structures—with a person's 'taste' or 'character' (BGE 260, 287). This does not mean, of course, that someone's actions are or should be simply irrelevant to our evaluative attitudes towards her. For there are compelling reasons for thinking—although Nietzsche himself does not

[20] For more detailed discussion, see Poellner (2007).

[21] See e.g. Lyons (1980: 8, 78–80); de Sousa (1987: 121–3); Mulligan (1998).

[22] Cf. Lyons (1980: 99–104); Nussbaum (2001: ch. 1); de Sousa (2002: 251); Doering and Peacocke (2002: 92–5).

stress this point—that the relation between a person's affective structure and her actions is not merely causal. 'Affects' such as the emotions of admiration or contempt in appropriate circumstances *motivate* actions which therefore can be said to express them, rather than being merely caused by them.[23] However, the question which concerns us here is whether Nietzsche's own affective evaluations of another's taste or character can be understood as having evaluative content which might be veridical. Can emotions like contempt veridically represent the value of *ressentiment*?

In trying to attain clarity on this issue, we need to bear in mind that the relevant concept of perception concerns the *phenomenal* features of the represented object—analogous to the colours as represented in sensory perception. Since emotions (purport to) represent value features of their objects, their representational content, on the perceptual model, would have to include the phenomenally intrinsic (dis)value of their object—in the present example, of *ressentiment*. The proper analysis of a conscious affective state such as Nietzsche's contempt for *ressentiment* therefore needs to exhibit a quite complex structure, since what that state is about is another intentional state, which can only be specified with reference to *its* objects. Now what I have called the phenomenally intrinsic disvalue of *ressentiment*, regarded as a conscious mental state, just is its experienced unattractiveness from the point of view of the subject who is in that state if the latter is adequately and transparently presented to the subject. *Ressentiment* being itself an intentional state, the subject's first-personal, transparent, and adequate awareness of it would of course have to include an explicit awareness of the relevant features of its intentional objects and of the aspects under which they are represented in it. What are these aspects? On Nietzsche's account, *ressentiment* is essentially a form of negative affect (hatred) towards its target, a hatred which the subject however does not avow as such, but rather interprets as *moral disapproval* (see esp. GM I 10). This reinterpretation of negative affect as moral righteousness typically involves, furthermore, a motivated distortion of the object's character (through selective attention, purposeful ignoring of countervailing evidence, etc.). An adequate, transparent awareness of *ressentiment* by the subject herself would therefore require explicit consciousness of all these essential constituents. It is plausible to claim that, if *ressentiment* were thus adequately presented to the subject herself, she could not *fail* to be aware of its intrinsic phenomenal disvalue or unattractiveness. The reason why she is normally not explicitly aware of the latter, and thus can unproblematically remain in that state, is precisely that the emotion is *not* presented adequately to her—that she is in a state of 'self-deception' (*Selbstbetrügerei*; GM I 13).[24] *Ressentiment* is only stable as

[23] For the relevant concept of motivation, see esp. Stein (1980: 56–60, 94–5). Also Wrathall (2005).

[24] For a detailed analysis of *ressentiment*, see Poellner (2004).

an emotion to the extent that it involves such self-deception—a motivatedly erroneous self-awareness.

There may be thought to be a tension between the characterization of phenomenal objectivity offered earlier and the concept of intrinsic phenomenal value. Isn't such value a paradigmatic example of a *subjective* property? No, for while it clearly is not independent of subjectivity (just as phenomenal colours aren't) it is not *purely* subjective either. In so far as it is based on a reference to intentional objects, it is not simply identical with any one experience or with a part or aspect of such an experience.[25] It is therefore experientially accessible to the subject by way of numerically distinct representations of these objects. Moreover, it is also accessible to *others* if a certain view of our 'knowledge of other minds' can be vindicated, according to which we can, in favourable circumstances, directly perceive (rather than infer) others' psychological states, such as *ressentiment*, through their expressive behaviour and actions. On this view, which is familiar from classical phenomenology and from some interpretations of the later Wittgenstein, the relation between a mental state and possible behavioural expressions of it is constitutive rather than merely causal.[26] If this sort of approach is right, then others' conscious mental states are, and perhaps must be, non-inferentially accessible not only to themselves but also to others. But it may still be doubted whether the intrinsic phenomenal *value* features of such states are also intersubjectively experientially accessible. One response to this doubt would be to say that the perception of another's mental state can non-inferentially motivate a *reproduction* of its evaluatively relevant aspects in a suitably sensitive observer. Thus, the phenomenal objectivity of these features would attach to them not *qua* particular instances but *qua* types that are multiply exemplifiable by virtue of perceptions of the same particular mental state by different observers. Another, in my view more attractive, solution would be to insist that the perception of the phenomenal value features of another's conscious mental state does not necessarily require a prior, temporally contiguous, reproduction or simulation of that state in oneself at all. Rather, affective responses such as revulsion or contempt sometimes just *are* third-personal perceptions of those value features, non-inferentially based on perceptions of verbal and other actions and behaviours expressing the state.

We can now see in what sense Nietzsche's contempt for *ressentiment* might be a veridical representation of that state. To the extent that *ressentiment* is transparently presented to the subject herself, it cannot fail to exhibit an intrinsic phenomenal disvalue, and the latter is just what is captured by revulsion at it. Appreciating this point also may help us see what qualifies Nietzsche's discussion

[25] An intentional object in Husserl's classical, broad definition is an item whose existence or nature is not exhausted by ('transcends') what is given in any one intentional experience. It is essentially something which is representable, without being itself modified, in numerically different intentional experiences. See Husserl (1973: § 13).

[26] See e.g. Stein (1980: 85–96); and McDowell (1982).

of *ressentiment* as a *criticism* of it. His basic charge—not his only one, to be sure—is that r*essentiment* is an intrinsically unattractive state to be in, if seen in the open. No one who accepts Nietzsche's account of it can either desire to be in it or wish to remain in it, except for instrumental reasons (although it is hard to see why *ressentiment* should be *indispensably* instrumental for any projects whose aims do not include it).[27] The example therefore also illustrates one sense in which Nietzsche's own ethical commitments are, as has often been noted, quite close to certain kinds of aesthetic evaluations. Both are appraisals of intrinsic phenomenal features of their objects.

But to accept that emotions are 'ways of seeing', that they are or include potentially veridical perceptions of value in the sense sketched above, also entails that they have conditions of success, which fail to be met when they are *non*-veridical. Nietzsche himself acknowledges this point in criticising *ressentiment*, in part, for *mis*representing both itself and its targets (GM I 10; GM III 14). His criticism implies, therefore, that if his own contempt for *ressentiment* misrepresented that condition's intrinsic phenomenal (dis)value, it would be inappropriate and in need of revision (or 'purification'; cf. GS 335). It has often been said, by critics and admirers alike, that the ethos Nietzsche seems to be attracted to is ultimately a solitary one and that there is no role, or at least no positive role, for intersubjectivity in it. If my argument so far is along the right lines, this is true only with significant qualifications. For if I am right, the ethos of the sovereign individual necessarily includes affective responses to the subjective states of others which aim at veridical access to those states—at a kind of *Verstehen*. Nietzsche privileges, to be sure, a cultivated attentional bias

[27] These claims may seem controversial. For Nietzsche says that *ressentiment* does have value for those subject to it—it makes their lives more tolerable than they would otherwise be by enabling them to transform the passive suffering of their condition into a form of activity (GM I 10). He also sometimes suggests that positive affective states are only *possible* relative to, and as 'overcomings' of, negative states (i.e. of 'suffering'; cf. WP 693, 695). As for the first of these points, it is possible, albeit far from clear, that for Nietzsche the *ressentiment* of the slaves has, relatively, more *intrinsic* phenomenal value than the self-contempt that would result from their adopting their masters' judgement of themselves (but see BGE 62 for doubts on this score). However, this still would not make *ressentiment* into a positive or desirable state any more than an amputation of a limb is desirable merely because it may be better than wasting away with gangrene. *Ressentiment* may of course have a certain *instrumenta*l value, e.g. in fostering 'cleverness' in its subjects (GM I 10). But even if the latter quality is seen as an unambiguously positive, Nietzsche nowhere suggests that *ressentiment* is *necessary* for its acquisition—had the slaves in his story attempted to outwit their masters in less rancorous ways, they presumably would have had no less need to be clever. As for the idea that all experience of positive value requires a contrast with, and overcoming of, displeasure or suffering (and hence experiences of intrinsic phenomenal disvalue), even if this is granted, it would be only a contingent fact about humans that suffering has instrumental worth in stimulating activity, which aims at overcoming it. Its 'instrumentality' would in fact simply *be* its undesirability. (A far more problematic claim, which Nietzsche sometimes seems to come close to, would be that there is a *conceptual* dependence of the experience of positive value on its opposite). In any case, one of Nietzsche's central charges against Christianity is precisely that by *disguising* suffering, e.g. through fostering *ressentiment*, it prevents suffering from being instrumental to the creation of positive value (BGE 225; cf. Leiter, 2002: 128–34).

towards 'affirmation', in other words, towards the *positive* values instantiated in some of those subjective states. Hence he extols, for instance, 'that enthusiastic impulsiveness in anger, love reverence, gratitude, and revenge by which noble souls have at all times recognized one another' (GM I 10).

But while certain aspects of the value-disclosing disposition that Nietzsche calls 'noble' may constitute (part of) his substantive ideal, does this have any relevance for the 'transcendental' question concerning full personhood and free agency? Cannot *ressentiment*-subjects in principle also be free agents, for all that has been said here? In so far as they understand their affective-evaluative life as essentially aiming at the truth (about themselves or others), even if in fact they fail in this, they would seem to meet the substantive constraint on Nietzschean freedom I have outlined. On the other hand, to the extent that *ressentiment* is a form of self-deception requiring a pre-reflective, yet motivated distortion of the subjectivity of another, it disavows the aim of veridicality at the pre-reflective level. It is partly for this reason that the *ressentiment*-subject *cannot identify* with her pre-reflective (distorting, slandering, etc.) motives and has to avoid acknowledging them, if she is to remain in that state. And to the extent that she cannot identify with the affects that actually dominate in her and determine her actions, she is unfree. This example might also suggest a more concrete understanding of the formal condition of 'wholeness' or unity-in-diversity that is necessary for Nietzschean freedom. As we saw in Section 1, not every kind of unity will meet Nietzsche's requirements. Any vertebrate organism is a highly unified system of diverse functions, but it is not autonomous by dint of that. Nor is a brainwashed individual with a complex system of bizarre preferences, however protracted they may be and however content he may be with all of them as a result of electrodes having been applied to the right areas of his brain. The account canvassed above suggests a possible solution to the problem of what constitutes the right kind of unity-in-diversity. The relevant *diversity* ('multiplicity of drives') might consist in a maximal openness to the phenomenal value features actually or potentially exemplified in the life-world. And the *unity* in question is perhaps most fruitfully understood as the ability to enact an orientation in one's life that reflects the internal ordering or 'rank' among these features. An orientation that systematically and for non-instrumental reasons preferred *ressentiment* to some of the character traits Nietzsche classifies as 'noble' (generosity, honesty towards oneself, etc.) would be an example of the wrong kind of unity.[28]

[28] The idea of an order of values, considered as phenomenal features of the life-world, need not in principle be more problematic than that of an order of colour shades in respect of lightness and darkness, notwithstanding the evident fact of widespread evaluative disagreement. Such disagreement by itself merely shows that individuals respond differently to the same *objects* (persons, etc.), not to those objects under the same aspects or *modes of presentation*. There is, to be sure, a question as to whether the idea could plausibly be applied to the putative values of unconscious items (such as art works or landscapes). A fuller discussion of these issues must wait for another occasion.

3. THE FREE SPIRIT AND THE CRITIQUE OF THE WILL TO TRUTH

Besides its role in the constitution of full-fledged, 'autonomous' personhood, freedom also figures as a substantive *evaluative ideal* in Nietzsche's later philosophy. In this concluding section, I want to focus on one component of his talk of freedom in this latter context that has received relatively little attention in the literature. We encounter it in passages like these:

> All these pale atheists, anti-Christians, immoralists, nihilists; . . .—they certainly believe that they are as completely liberated from the ascetic ideal as possible, these 'free, *very* free spirits'; and yet, to disclose to them what they themselves cannot see—for they are too close to it themselves: this ideal is precisely *their* ideal, too . . .—if I have guessed any riddles, I wish that this proposition might show it!—They are far from being *free* spirits: *for they still have faith in truth*. (GM III 24)

I do not think that we do full justice to Nietzsche's point here, and to the revolutionary significance he associates with it, if we take him to be saying merely that there may be more important things than either purely theoretical or scientific enquiry—after all, many (probably most) people who are not professional philosophers or scientists would be willing to agree to that. I submit that in order really to take seriously the radicality Nietzsche himself sees in his position, we should take him to be proposing a quite novel way of interpreting the relation between our practical, necessarily evaluating, 'life', and our theoretical, and in particular our metaphysical, beliefs. One traditional way of interpreting this relation is unambiguously rejected by him, although this rejection is hardly original. As the passage from GS 301 cited earlier indicates, the *phenomenal* objectivity of many values often gives rise to the delusion that values are *metaphysically* objective: that they are existentially and/or conceptually independent not just of any particular affective experience, but of such experiences altogether. The temptation is to think of affective experiences as *discovering* value properties exemplifiable without them, rather than as *co-constituting* value. Nietzsche emphatically denies that values are metaphysically objective properties:

> Whatever has *value* in our world now does not have value in itself, according to its nature—nature is always valueless—but has been given value at some time, as a gift; and *we* were the givers and bestowers! Only we have created the world *that concerns human beings*! (GS 301)

> Truly, human beings have given themselves all their good and evil. Truly, they did not take it, they did not find it, it did not descend to them as a voice from heaven. Humans first placed values in things. . . . Only through valuating is there value. (Z, 'Of the Thousand and One Goals')

According to Nietzsche, value existentially depends on the existence of affective experiences and is actualized only in or through them. This is what he means when he says that value is not 'found' but 'created' by 'evaluating' (GS 301), that morality is only a sign-language of the affects (BGE 187), and that '*every* ideal presupposes *love* and *hatred, admiration* and *contempt*' (KGW VIII.2.10.[9]). And many of his formulations (e.g. Z, 'Of the Way of the Creator') suggest that achieved freedom in his sense essentially involves the *acknowledgement* of this truth.

It is worth reflecting briefly on the prospects of a radical *revision* of the ordinary concept of value which would construe 'value' as independent of affective phenomenal properties, analogously to revisions of our 'folk' concepts of the colours, say, as disjunctions of non-dispositional microphysical properties or as light dispositions. While it is at least conceivable that primary-quality concepts of the colours might largely—but not entirely—replace the phenomenal folk concepts of them, although such a replacement would be enormously impractical in many contexts, it is simply not credible to suppose that the proposal of an analogous revision of the everyday generic concept of value might hold out any prospect of success. For in order to be successful, humans would need to be persuaded that what should guide their actions should be completely independent of their (affective) experience of value.[29] But what should possibly persuade them of *that*? For conceptual revisions to hold any attractions for us, the latter must themselves by recognizable by us. But, Nietzsche contends, we are not only attached to the specific goods we value or desire, but more fundamentally to valuing and desiring, as we currently understand these modes of engagement with the world, themselves. This is the crux of Zarathustra's dictum that 'valuating is itself the value and jewel of all valued things' (Z, 'Of a Thousand and One Goals'), and of Nietzsche's remark in his own voice that what matters most deeply to us is not what is desired in any given instance, but the *desiring itself* (KGW VII.1.20.[4]; p. 661). It is because we have a non-negotiable interest in how things consciously *matter* to ourselves and others—because we profoundly *value* the idea of intelligibility as consciously authenticable *significance*—that the affective phenomenal properties figuring in many ordinary psychological explanations have a central and irreplaceable role in our cognitive lives.[30] Even the ascetic ideal, for all its deficiencies, offers a kind of significance. Wanting

[29] Note that such a stance would be much more radical than Kant's, who conceded that moral value is experientially accessible through the 'intellectual' emotion he called *respect* (*Achtung*).

[30] According to the Nietzschean transcendental argument sketched at the beginning of Section 2, abandoning the experience-involving concepts of desire and value would entail ceasing to think of ourselves as subjects (and a fortiori as free, properly self-directing subjects). This point is also clearly present, allowance being made for the rhetorical hyperbole, in the following passage from *Zarathustra*: 'And do you tell me, friends, that there can be no dispute over taste and tasting? But all life is dispute over taste and tasting! Taste: that is at the same time weight and scales and weigher; and woe to anything living that wants to live without dispute over weight and scales and weigher!' (Z, 'Of the Sublime Men').

to do away with 'sense', or conscious significance, altogether, would amount to wanting to abolish humanity:

> . . . *the ascetic ideal offered man meaning!*. . . . man was *saved* thereby, he possessed a meaning, he was no longer like a sheaf in the wind, a plaything of nonsense—the 'sense-less'—he could now *will* something; no matter a first to what end, why, with what he willed; the *will itself* was saved. (GM III 28)

Nietzsche's position as I have explicated it so far conjoins a *phenomenal objectivism* about value (or about many values) with a *metaphysical anti-objectivism*. Even if an absolute conception of the world made sense, values would not figure in it. This view about the relation between the world as it shows up for us in our practical engagement with it—our life-world—and the world as it is in itself, assuming the latter conception to have content, is unusual, although it is not unique to Nietzsche. In fact, however, his actual position on the relation between evaluative practice and metaphysical belief is more complex and radical. This becomes evident in the light of statements like the following: 'The falseness of a judgement is to us not necessarily an objection to a judgement: it is here that our new language perhaps sounds strangest. The question is to what extent it is life-advancing . . .' (BGE 4). Elsewhere he adds: 'It is of cardinal importance that one should abolish the *real world*. It is the great inspirer of doubt and devaluator in respect of the world *we are*' (WP 583; cf. BGE 34; TI V). I take it that by 'abolishing the real world' Nietzsche means: ceasing to grant the conception of a metaphysically real world which in principle might be quite different from our life-world in terms of its essential properties any fundamental or overriding importance in our cognitive and evaluative economy.[31] The idea of such a metaphysically real world could of course only inspire doubt and devalue the 'world we are' if it could potentially produce a conflict between what we are committed to at the level of practical 'life' and our metaphysical beliefs. If the contents of our practical commitments were simply insulated from the answers to metaphysical questions, then no such answers could 'inspire doubt' about 'the world we are' in the first place, and the need for 'abolishing the real world' would hardly be pressing. It is therefore clear that Nietzsche envisages at least the possibility of a *theoretical* conflict between the level of evaluative practice and the level of metaphysical belief.[32] What might such a conflict look like? Well, it might turn out that an adequate analysis of our aesthetic practices would show them to involve the belief that phenomenal value properties of objects—the beauty of a painting, say—have *causal powers*: that they genuinely *affect* us. Or various ethical practices might cease to be intelligible to ourselves unless we continued

[31] Note that in none of the four passages referred to does Nietzsche contend that the idea of the world having some ultimate, intrinsic nature is simply inapplicable or *incoherent*.

[32] In this respect, too, Nietzsche's approach differs from views according to which the contents of our practices are, rightly understood, logically independent of metaphysics. For the latter position, see e.g. Blackburn (1985: 3–4).

to think that certain phenomenal, and essentially value-involving, psychological properties—such as an experience's *painfulness*—themselves caused (as well as justified) us to act in certain ways. The point is not that these beliefs, for all we currently know, are essential to the relevant practices, but that they might turn out to be so. And if they do, then there will be a theoretical conflict between those practices and certain metaphysical propositions we might also have good reasons to subscribe to—such as the idea that neither experienced pain nor aesthetic beauty *really* have causal powers.

Nietzsche's claim, in BGE 4, that 'the falseness of a judgement is to us not necessarily an objection to the judgement', is most plausibly interpreted as the enjoinder that, if such a theoretical conflict between practice and metaphysical belief were to arise, we should affirm the former over the latter, provided that we have strong independent reasons to value the practice (e.g. on account of its being 'life-enhancing'): 'Why could the world *which is of any concern to us*—not be a fiction?' (BGE 34). This seems to me to be the very heart of the 'critique' of the 'will to truth' that Nietzsche calls for in his later writings (GM III 24). And the attitude that has just been described is one of the most central aspects of his own evaluative ideal of genuine '*freedom* of spirit' (ibid.).

It is important to recognize that this attitude is not a form of irrationalism. The free spirit who is unbeholden to the will to truth in Nietzsche's sense certainly recognizes truth-based reasons for judgement and action. It is only a certain kind of metaphysical truth—call it *purely theoretical* truth—which, assuming it to be accessible to human enquiry, the free individual envisaged by Nietzsche would disregard in practical contexts wherever such truth conflicted with beliefs implicated in his deep practical commitments. To get a grip on what such purely theoretical truths might be, it helps to approach them by way of a contrast with truths which are not purely theoretical, such as scientific truths. The reason why the experimental laws and theories of science are not purely theoretical is that they have predictive consequences. If a physical theory about a certain domain of phenomena is true, then it will be suited to yield, in conjunction with appropriate statements of initial conditions, correct predictions about phenomena within that domain. In so far as scientific theories predict the future course of an appropriately placed attentive observer's experience, they evidently have, or are essentially apt to have, practical relevance. But neither the results nor the methodology of the natural sciences are necessarily in conflict with any practical commitments which could plausibly be claimed to be *fundamental* to our human evaluative—ethical or aesthetic—practices, even if (as I mooted earlier) these commitments should include the belief that our conscious responses to the world can be affected by value-involving phenomenal properties like a friend's pain or a sculpture's beauty. Now, it may be protested that this belief, and hence any practice which implies it, *is* incompatible with the methodological principles of a properly scientific psychology, according to which human behaviour is to be explained with recourse to the properties and entities recognized by the natural sciences (broadly

conceived), which of course do not include phenomenal properties like values. But this objection rests on a misunderstanding. What is constitutive for a scientific psychology is the following assumption: the properties recognized by the physical sciences (call them physical properties) are sufficient to account nomologically for psychological properties and human behaviours, which can all be considered as elements of the physical world; and apparent phenomenal properties such as values have no autonomous causal powers, that is, powers which could vary independently of physical states, or causally affect physical states, or both. The physical, including the behaviour of the human organism, is in this sense a closed causal system. But this assumption, which is all that is required for naturalistic psychology as a prediction-generating science (rather than as non-predictive metaphysical theory), is entirely compatible with accepting both the *indispensability* and the *truth* of the phenomenological explanations essentially pertaining to everyday evaluative practices—explanations such as 'he chose to buy a recording of Strauss's *Four Last Songs* because he was moved by their beauty when he heard them on the radio'. The methodological principles of scientific psychology and the commitments involved in everyday practices and explanations square with each other just in case (a) the latter give us information, including rough generalizations sustaining true counterfactuals—'he wouldn't have bought the recording if he hadn't been moved by the beauty of the music'—information which is itself prized by us and which is not made available by naturalistic types of explanation;[33] and (b) the phenomenal properties referred to are strongly supervenient on (co-variant with) physical properties simultaneously exemplified.[34]

What is needed for the project of a scientific psychology is thus no more than the strong supervenience of phenomenal psychological properties on physical properties. But supervenience, it is often said, stands itself in need of explanation. Why do phenomenal and physical properties co-vary nomologically, if they do?[35] This demand for a further 'explanation', once the basic level of co-variation is reached, is not itself internal to science, since no predictions could possibly be relevant to settling the issue. It is, rather, a metaphysical demand. And it is only once this demand is conceded, and certain explanatory answers are in the offing, that there may emerge an inconsistency between these proposed extra-scientific, metaphysical explanations of nomological co-variance and beliefs essentially implicated in any evaluative practice we can make sense of.[36]

[33] As I argued earlier, the fact that such information has non-negotiable, overriding importance for us is just what Nietzsche insists on when he makes Zarathustra say that 'valuing is itself the value and jewel of all valued things' (Z, 'Of a Thousand and One Goals').

[34] The relevant notion of strong supervenience I have in mind is Kim's (1993: 65).

[35] See Kim (1993: 144–8); Shafer-Landau (2003: 77–8). The feeling that even a strong supervenience relation is not 'really' or 'ultimately' explanatory is connected with the fact that it is weaker than the relation of dependence. On this, see Charles (1992: 272–5); and Kim (ibid.).

[36] While this way of balancing the legitimate claims of scientific naturalism and phenomenology may be attractive, is there any direct evidence that Nietzsche is sympathetic to it? Well, he evidently

So, when Nietzsche urges a 'critique' of the will to truth, and refusing to be agitated about the prospect that the practical 'world which is of any concern to us' might be a 'fiction', is he suggesting that deep practical commitments should simply override *any* true metaphysical beliefs that might, if granted importance, undermine them? That this cannot be quite the right account of the matter becomes clear once we recall that Nietzsche's own attitude towards metaphysics is not simply globally indifferent or quietist. On the contrary, he spends considerable effort on criticizing certain kinds of metaphysics, above all Christian metaphysics, for their theoretical errors, which therefore are presumably seen by him as potentially mattering (see e.g. A 15). The reason why he considers it important not simply to ignore, but to reject such metaphysics, is that they do not only comprise *purely* theoretical propositions in the sense defined earlier—propositions that have no predictive consequences. Christian metaphysics certainly does have such consequences. If it is true, then we will all have a conscious life after this present earthly life which is consciously continuous with the present life. We will be aware of our former earthly lives in that future life and will think of ourselves as numerically the same person who lived that earthly life. And according to the mainstream versions of Christianity, our present choices and actions will very much affect the nature of that future life and we will then be in a position to recognize this. The question whether Christian metaphysics is true therefore has an evident practical, experiential relevance for us irrespective of whether we are interested in knowing the truth for its own sake. It is this practical dimension which gives Christian metaphysics its soteriological significance, dramatized in Pascal's Wager. Pascal's point is precisely that, for *practical* reasons, we cannot be simply indifferent with respect to the question whether Christianity is true.

While certain metaphysical theories thus have an evident and incontestable practical dimension, there are others that do not. Among these purely theoretical claims are many that historical and contemporary metaphysicians have been, and continue to be, preoccupied by. They include the following familiar staples: that the real spatial world consists of absolute, non-perspectival objects (Descartes, Locke); that there are no non-perspectival spatial objects (Berkeley, Kant,

gives a central place, not only in his own practice, but also in his second-order reflections on it, to phenomenological psychology involving reference to 'valuations': 'our opinions, valuations, and tables of the good certainly belong among the most powerful levers in the mechanism of our actions, but in each individual case the law of their mechanism is indemonstrable' (GS 335). The most natural way of reading his point here is that in standard cases, personal-level phenomenological explanations are *indispensable* for satisfying our cognitive interests, although they are not *sufficient* explanations, since the personal level where we can properly speak of 'actions' is necessarily subtended by a subpersonal 'mechanism' the details of which are largely unknown. Nietzsche also suggests that, in certain non-standard cases, which according to him are more frequent than we are inclined to think, personal-level explanations are inappropriate and should ideally be replaced by what he calls 'physiological' explanations, presumably analogous to clinical practice in diagnosing and treating certain cases of depression. (See GM III 17; TI VI 6).

sometimes Nietzsche); that the world does not include phenomenal consciousness among its ultimate realities (physicalism); that the real world consists only of consciousness and its objects (idealism); or that being is at the fundamental level in principle non-representable (Heidegger, sometimes Nietzsche). They also include the various metaphysical explanations of the apparent supervenience of phenomenal upon physical properties. One such explanation might be that phenomenal properties are higher-order properties of sets of physical properties and that phenomenal objects are token-identical with sets of physical items. More outlandish alternative explanations might invoke a metaphysical parallelism of the Spinozist type, or some form of occasionalism or idealism. Yet another instance of purely theoretical questions, famously bypassed by Newton and the physical sciences he inaugurated, concerns the intrinsic qualitative nature of actualizations of those real forces whose numerical indices figure in the equations of physics.

What makes all these questions and theories *purely* theoretical is that none of them constitutively generate predictive consequences and their only possible warrant is therefore a priori[37] — unless they are supplemented by more substantive doctrines, for example about the existence or non-existence of immortal persons, or of divine providential plans, or the possibility or impossibility of an afterlife, and so forth. Without such supplements, it makes no experiential difference whether the token-identity theory or metaphysical parallelism, physicalism or idealism, panentheism or the conclusion of the cosmological proof, is true or even known to be true. Or rather, such knowledge can only be of experiential significance for us if we happen to have *autonomously belief-dependent interests*; if, that is, our well-being can be affected by knowledge or belief about an empirically, and hence practically, inaccessible truth. (A person with such autonomously belief-dependent interests — the 'will to truth' castigated by Nietzsche — might, for example be depressed or buoyed up by the belief that being is 'in itself' in principle unrepresentable.)

But can't purely theoretical metaphysical beliefs rationally affect certain practical interests at least *negatively*? If someone lives in the hope of an Elysian afterlife for her immortal soul, then this hope would seem to be dashed by her acceptance of an a priori argument to the effect that *really* there are only material particles in motion, or fields of non-conscious energy. The response suggested by Nietzsche's remarks in BGE 4 is that, *if* she had good practical reasons to regard her commitments as 'life-enhancing', then she should only be troubled by such a priori counter-arguments if the latter could also show that the actual world is *in fact* unable to sustain or produce a post-mortal,

[37] I count among a priori arguments also so-called inferences to the best explanation, wherever the good-making features of the explanation proffered do not include superior predictive power.

beatific, personal consciousness (however metaphysically 'fictional' its contents might be—remember BGE 34). But how could the metaphysical theory by itself show that? The work of disenchantment here is not done by the physicalist metaphysics, which is in fact quite otiose, but by its supplementation with certain empirical beliefs quite independent of it, and shared by many, about consciousness and the limitations it may be subject to, given what we think we know about its actual supervenience base.

In the light of these reflections, I think it is possible to circumscribe more precisely in what sense the Nietzschean free spirit is not subject to the will to truth. The really new and 'strangest' point suggested by his remark that 'the falseness of a judgement is to us not necessarily an objection to a judgement' (BGE 4), and by his rhetorical question why the 'world which is of any concern to us' should not be a 'fiction' (BGE 34), is this. The free spirit who represents Nietzsche's evaluative ideal has no significant autonomously belief-dependent interests or values. As a consequence, even if beliefs about the phenomenal world implied by her deepest practical commitments were logically incompatible with purely theoretical metaphysical truths, provable on a priori grounds (and on such grounds alone), she would refuse to revise her practical commitments on the basis of those grounds. She would have 'abolished' the real world in accepting no other grounds for her orientation to the world than experientially accessible, in the last resort affectively authenticated, reasons neither in need of metaphysical underpinning, nor subject to metaphysical refutation. Nietzsche's 'free spirits *par excellence*' (GM III 24) are indifferent as to whether the life-world which engages their practical concerns might turn out be 'fictional' in the light of purely theoretical truths established by a priori metaphysical enquiry.[38] The latter do not constrain the free spirit's practice, and it is this fact which is expressed rhetorically by Nietzsche in citing the slogan 'nothing is true, everything is permitted' (ibid.).

What I hope to have shown in this paper is that Nietzsche's account of freedom and the free person cannot be detached from the question of an individual's relation to truth. As we have seen, his analysis of this relation and its relevance to freedom is complex and nuanced. I have suggested that an affective orientation towards phenomenally objective value is, according to him, a necessary constituent of full, 'autonomous' personhood. Yet, indifference towards the question of the metaphysical reality of what is thus valued is a central component of the kind of freedom that figures as an essential part of Nietzsche's substantive ideal.[39]

[38] Nietzsche himself of course often expresses doubts about the availability of such metaphysical knowledge. For discussion of this sceptical strand in his later thought, see Poellner (1995: 57–78, 150–61, 173–86).

[39] For helpful comments on an earlier version of this essay, thanks are due to the participants of the Nietzsche workshop organized by Ken Gemes at Birkbeck College, London, in April 2006.

REFERENCES

WRITINGS BY NIETZSCHE

[1878–80] *Human, All-too-Human*, trans. R. J. Hollingdale (Cambridge: Cambridge University Press. 1986).
[1882–87] *The Gay Science*, trans. Walter Kaufmann (New York: Vintage, 1974).
[1883–85] *Thus Spoke Zarathustra*, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1969).
[1886] *Beyond Good and Evil*, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1982).
[1887] *On the Genealogy of Morals*, in *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage, 1989).
[1888] *The Case of Wagner*, in *The Birth of Tragedy and the Case of Wagner*, trans. Walter Kaufmann (New York: Vintage, 1967).
[1889] *Twilight of the Idols*, in *Twilight of the Idols and The Anti-Christ*, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1990).
[1895] *The Anti-Christ*, in *Twilight of the Idols and The Anti-Christ*, trans. R. J. Hollingdale (Harmondsworth: Penguin, 1990).
[1906] *The Will to Power*, trans. Walter Kaufmann and R. J. Hollingdale, ed. Walter Kaufmann (New York: Vintage, 1968).
[1908] *Ecce Homo*, in *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage, 1989).
Werke: Kritische Gesamtausgabe. Giorgio Colli and Mazzino Montinari, et al. (eds.) (Berlin: de Gruyter, 1967–).
Briefwechsel: Kritische Gesamtausgabe. Giorgio Colli, Mazzino Montinari et al. (eds.) (Berlin: de Gruyter, 1975–2004).

OTHER WORKS CITED

Blackburn, Simon (1981). 'Rule-Following and Moral Realism', in S. Holtzmann and C. Leich (eds.), *Wittgenstein: To Follow a Rule* (London: Routledge).
____ (1985). 'Errors and the Phenomenology of Value', in T. Honderich (ed.), *Morality and Objectivity* (London: Routledge).
Charles, David (1992). 'Supervenience, Composition, and Physicalism', in David Charles and Kathleen Lennon (eds.), *Reduction, Explanation, and Realism* (Oxford: Clarendon Press).
Dancy, Jonathan (1993). *Moral Reasons* (Oxford: Blackwell).
De Sousa, Ronald (1987). *The Rationality of Emotion* (Cambridge, Mass.: MIT Press).
____ (2002). 'Emotional Truth', *Proceedings of the Aristotelian Society*, supp. vol. 76: 247–75.
Doering, S., and Peacocke, Christopher (2002). 'Handlungen, Gründe und Emotionen', in S. Doering and V. Mayer (eds.), *Die Moralität der Gefühle* (Berlin: Akademie Verlag).
Gemes, Ken (2009). 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual' (this volume).
Guay, Robert (2002). 'Nietzsche on Freedom', *European Journal of Philosophy* 10: 302–27.

Husserl, Edmund ([1938]1973) *Experience and Judgement* (Evanston: Northwestern University Press).
Kant, Immanuel ([1781–87]1950). *Critique of Pure Reason*, trans Norman Kemp Smith (London: Macmillan).
Kim, Jaegwon (1993). *Supervenience and Mind* (Cambridge: Cambridge University Press).
Leiter, Brian. (2002). *Nietzsche on Morality* (London: Routledge).
Lyons, William (1980). *Emotion* (Cambridge: Cambridge University Press).
McDowell, John (1982). 'Criteria, Defeasibility and Knowledge', *Proceedings of the British Academy* 68.
____ (1985). 'Values and Secondary Qualities', in Ted Honderich (ed.), *Morality and Objectivity* (London: Routledge).
Mulligan, Kevin (1998). 'From Appropriate Emotions to Values', *Monist* 81: 161–88.
Musil, Robert ([1930–33]1995). *The Man without Qualities*, 3 vols. (London: Minerva).
Nehamas, Alexander (2001). ' "How One Becomes What One Is" ', in J. Richardson and B. Leiter (eds.), *Nietzsche* (Oxford: Oxford University Press).
Nussbaum, Martha C. (2001). *Upheavals of Thought: The Intelligence of Emotions* (Cambridge: Cambridge University Press).
Pears, David (1985). *Motivated Irrationality* (Oxford: Oxford University Press).
Poellner, Peter (1995). *Nietzsche and Metaphysics* (Oxford: Clarendon Press).
____ (2004). 'Self-Deception, Consciousness and Value: The Nietzschean Contribution', *Journal of Consciousness Studies* 11 (10–11): 44–65.
____ (2007). 'Affect, Value and Objectivity', in Brian Leiter and Neil Sinhababu (eds.), *Nietzsche and Morality* (Oxford: Oxford University Press).
Scheler, Max ([1913–16]1973). *Formalism in Ethics and Non-Formal Ethics of Value* (Evanston: Northwestern University Press).
Shafer-Landau, Russ (2003). *Moral Realism* (Oxford: Oxford University Press).
Stein, E. ([1917]1980). *Zum Problem der Einfühlung* (Munich: Kaffke).
Strawson, Galen. (1994). *Mental Reality* (Cambridge, Mass.: MIT Press).
Taylor, Charles. (1982). 'Responsibility for Self', in G. Watson (ed.), *Free Will* (Oxford: Oxford University Press).
____ (1989). *Sources of the Self* (Cambridge: Cambridge University Press).
Watson, Gary. (1982). 'Free Agency', in G. Watson (ed.), *Free Will* (Oxford: Oxford University Press).
____ (2004). *Agency and Answerability* (Oxford: Clarendon Press).
Wrathall, Mark (2005). 'Motives, Reasons, and Causes', in T. Carman and M. B. N. Hansen (eds.), *The Cambridge Companion to Merleau-Ponty* (Cambridge: Cambridge University Press).

This page intentionally left blank

9

Nietzsche's Intentions: What the Sovereign Individual Promises

Aaron Ridley

> Signs of nobility: never thinking of degrading our duties into duties for everybody; not wanting to delegate, to share, one's own responsibility; counting one's privileges and their exercise among one's *duties*.
>
> *Beyond Good and Evil* 272

The 'sovereign individual' only makes one explicit appearance in Nietzsche's works, at the beginning of the second essay of the *Genealogy*. He is, we are told, 'a ripe fruit, but also a *late* fruit' (GM II 3); he stands at the 'end [of a] tremendous process' (GM II 2) during which man was made '*calculable, regular, necessary*, even in his own image of himself' (GM II 1). The process itself was the 'labour' of the 'morality of mores'—'the labour performed by man upon himself during the greater part of the existence of the human race, his entire *prehistoric* labour'; and 'notwithstanding the severity, tyranny, stupidity and idiocy involved in it', it was through 'the morality of mores and the social straitjacket' that 'man was finally *made* calculable'. And all of this was itself only preparatory; for 'when the tree at last brings forth fruit',

> we discover that the ripest fruit is the *sovereign individual*, like only to himself, liberated again from the morality of custom, autonomous and supramoral (for 'autonomous' and 'moral' are mutually exclusive), in short, the man who has his own independent, protracted will and the *right to make promises* . . . [Such a man] is bound to reserve a kick for the feeble windbags who promise without the right to do so and . . . who break [their] word even at the moment [they] utter it. (GM II 2)

The sovereign individual thus represents an achievement of some quite striking sort. He also raises a variety of questions and issues. One of these is where, in Nietzsche's genealogical narrative, he is supposed to be placed: is he prior to traditional (slave, Christian) morality, as some have suggested, or does he represent a recovery from that morality, as others have thought? Or does he (can he) occupy both positions? The text is inconclusive; and here I assume only that

he stands outside of traditional morality, whether he becomes before it or after it. This is an assumption that I return to in the final section of this essay. My first concern, however, lies with the sovereign individual's promise-making. For this, on the face of it at least, is puzzling—it is puzzling, that is, that his '*right*' to make promises should represent any sort of distinctive achievement.

I

According to Nietzsche's account, the 'prehistoric' effect of the 'morality of mores and the social straitjacket' was to produce a 'real *memory of the will*' (GM II 1): through the '*mnemotechnics*' of punishment man 'finally remembers five or six "I will not"s, in regard to which [he] had given [his] *promise* so as to participate in the advantages of society' (GM II 3). Man thus became '*calculable, regular, necessary*, even in his own image of himself'—and did so by making promises. And this, Nietzsche indicates, is true of socialized man *as such*. It is not true just of some sub-set of socialized man (to which the sovereign individual might be supposed to belong), but of everyone who has lived under the 'morality of mores'. And this would seem to rule out Nietzsche's 'feeble windbags' as an interesting possibility: these are socialized beings, after all—they know how to make promises; something counts as their 'breaking' their 'word'. So how can they not have the '*right*' to make promises (even if they fail—sometimes? often?—to honour the promises that they have made)? From this point of view, it seems that every socialized person must be a sovereign individual.[1]

But Nietzsche clearly does not mean that. He clearly thinks of the sovereign individual as something *more* than a merely socialized being with a handful of self-denying ordinances: the sovereign individual is, after all, 'liberated again from the morality of custom' (i.e. from just such a structure of 'I will not's). And the sort of achievement that he represents is brought out forcefully when one recalls a later description that Nietzsche offers of man's first imprisonment in the 'social straitjacket': 'the most fundamental change' that man ever experienced was that which

> occurred when he found himself finally enclosed within the walls of society and peace. The situation that faced sea animals when they were compelled to become land animals or perish was the same as that which faced these semi-animals, well adapted to the wilderness, to war, to prowling, to adventure: suddenly all their instincts were disvalued and 'suspended'. From now on they had to walk on their feet and 'bear themselves' whereas hitherto they had been borne by the water: a dreadful heaviness lay upon them . . . [I]n this new world they no longer possessed their former guides, their regulating, unconscious and infallible drives: they were reduced to thinking, inferring, reckoning . . . , these unfortunate creatures; they were reduced to their 'consciousness', their weakest and most

[1] A position of this sort is defended, I think, by Randall Havas (2000).

fallible organ! I believe there has never been such a feeling of misery on earth, such a leaden discomfort. (GM II 16)

The contrast with the sovereign individual could not be more striking: in his case, having 'the *right to make promises*', there is 'a proud consciousness, quivering in every muscle, of *what* has at length been achieved and become flesh in him, a consciousness of his own power and freedom, a sensation of mankind come to completion'; and this consciousness has 'penetrated to the profoundest depths and become instinct, the dominating instinct' (GM II 2). The physicalism of the imagery, and the point of that, is unambiguous: in the sovereign individual, culture has, as one might put it, become nature. He is borne up rather than oppressed by his socialization, and his capacities and dispositions are once again in harmony—as the sea animals' were—with the conditions of his existence. And he is *unusual* in this: to be as he is represents an achievement. As Nietzsche puts it elsewhere: 'high spirituality itself exists only as the ultimate product of moral qualities; . . . it is a synthesis of all those states which are attributed to "merely moral" men, after they have been acquired singly through long discipline and exercise, perhaps through whole chains of generations' (BGE 219).

Even the feeble windbag, we have to suppose, must be capable of promising in some rudimentary sense—to meet his buddies for a drink, say—and of seeing the promise through. This is the mark of someone who is socialized *at all*. But the sovereign individual is, as it were, socialized all the way down; and his special kind of '*right*' to make promises must be somehow connected to that (as, presumably, must be the kinds of promises that he makes).

II

When one promises in good faith to do something, one must, at a minimum, intend to do that thing. When it is true of the feeble windbag that he promises to meet his buddies for a drink, even he must intend to do so. This encourages the thought that, when the intention is in fact executed, the intention is responsible for the execution. And this isn't wrong. But it encourages another thought too: that the intention (perhaps construed as a discrete mental item of some sort) *causes* the action that constitutes its execution. And this thought *is* wrong—or so Nietzsche suggests.

In a much-discussed passage, Nietzsche claims that 'there is no "being" behind doing . . . —the deed is everything' (GM I 13).[2] This claim can be taken in a wider or a more restricted sense. At its widest—which is the sense that seems to interest Nietzsche most in the passage in question—the claim amounts to this: that what one does is nothing more (or less) than the direct expression of one's

[2] For an especially good discussion of this passage, to which I am indebted, see Pippin (2004).

nature. The strong act strongly, the weak act weakly. There is, as Nietzsche puts it, no 'neutral substratum' *behind* the doing, such that the 'strong' might instead have acted weakly, or the 'weak' strongly (ibid.). No: one *is* what one *does*. But Nietzsche can also be understood as advancing a more restricted claim, upon which any minimally interesting or plausible version of the wider claim must depend: that an agent's inner life cannot be understood *except* through what he does. To say that 'there is no "being" behind doing', in this restricted sense, is to deny that an agent's inner life is a self-contained realm of mental 'items'—of beliefs, desires, intentions, etc.—that stands in a merely contingent relation to (that stands 'behind') the agent's actions, so that it might, for instance, be said to *cause* them. And to say that 'the deed is everything' is to say that there is no inner reality that is ultimately independent of the outer—that what one does is not just not caused by, but is *constitutive* of, what one's beliefs, desires, intentions etc. are. The restricted version of Nietzsche's claim is therefore an expression of a generally anti-Cartesian picture of the nature of agency. The wider version is a radicalization of this.

The wider version is plausible, I think (it is, after all, the presupposition underlying any account of the virtues). But my argument requires only the restricted version, so in what follows I confine myself to that. I attribute to Nietzsche, that is, only the claim that what an agent believes, desires, intends and so on is not independent of what he does. The thought, say, that Ralph intends to get home on time is falsified if, *ceteris paribus*, he neglects the opportunity to do so when it is available to him. And Ralph himself, had he professed that intention, would, if he is honest, admit that he hadn't been entirely sincere, if only with himself. He didn't *really* intend to get home on time: at most, as it turned out, he hoped or wished or had it in mind that he might. The 'deed'—or, as in this case, the lack of it—'is everything'. And if Ralph had *promised* to get home on time, he has outed himself as a feeble windbag.

It is important to understand the force of the *ceteris paribus* clause in this. It does not necessarily count against someone's having intended to do something that they do not do it (even when the opportunity is available to them). For instance, if the reason why Ralph didn't get home on time, even though he could have done, is that a friend required urgent help, then other things weren't equal; and his failure to be punctual does not show that he didn't intend to be punctual. And similarly if Ralph had *promised* to be punctual. Because—when the *ceteris paribus* clause doesn't hold—his failure to get home on time doesn't show that he didn't intend to, it doesn't show, either, that his promise wasn't made in good faith. So if it is essential to a promise's being made in good faith that the agent intend to act on it, it is essential, too, that—*ceteris paribus*—he does indeed so act. This is a perfectly general point about intending and promise-making; and it must hold, if we are to take Nietzsche's remarks about the 'social straitjacket' seriously, of the feeble windbag just as much as of anyone else.

III

It is tempting, in light of this, to suppose that the sovereign individual must simply be someone who is especially fastidious about his *ceteris paribus* clauses—who, when he does promise something, gives as few hostages to fortune as possible. And some of the things that Nietzsche says certainly appear to encourage such a reading.[3] So, for instance, he notes that 'between the original "I will", "I shall do this" and the actual discharge of the will, its *act*, a world of strange new things, circumstances, even acts of will may be interposed' (GM II 1); and this might encourage the thought that it is, as it were, the frailty of our undertakings—given the contingencies of life—that prompts the sovereign individual to promise, as Nietzsche puts it, 'reluctantly, rarely, slowly' (GM II 2). On this construal, then, the sovereign individual is simply very very prudent: he promises only when he is extremely sure that other things will indeed remain equal, so that he really can 'stand security *for his own future*, which is what one who promises does!' (GM II 1). And if this were right, the feeble windbag would then just be someone who is a good deal less scrupulous in this regard.

There is, as we will see in a moment, at least *something* not too wide of the mark about this construal. But it is highly misleading, even so. For one thing, it threatens to bring the sovereign individual into perilously close alignment with the 'oppressed, downtrodden, outraged', who exhibit a 'prudence of the lowest order which even insects possess'. They say: ' "we weak ones are, after all, weak; it would be good if we did nothing *for which we are not strong enough*" ' (GM I 13)—for example, or so one might think, promising something when we're not quite sure that other things will be or remain equal. And it can't be right that the sovereign individual is simply someone who, doubting his own strength, has become a monster of caginess. The rhetoric of what Nietzsche says about him rules this out. But so, too, does the content. The sovereign individual is not only conscious of 'his own power and freedom', Nietzsche tells us, but is aware of how his 'mastery over himself also necessarily gives him mastery over circumstances, over nature . . . and over fate' (GM II 2), something that the 'oppressed' and 'downtrodden' manifestly lack, and which is the polar opposite, surely, of anything remotely to be described as 'prudence' (insect-like or otherwise).

Any account of the sovereign individual's '*right*' to make promises, then, must offer to make sense of his 'mastery' over 'circumstances', 'nature', and 'fate'. For it is this—or so the considerations offered above would suggest—that most plausibly marks what Nietzsche takes to be the *real* difference between

[3] —as, indeed, does the very image of a 'sovereign'. A sovereign had better, by and large, enter into his promises with a good deal of circumspection: much, after all, may hang on them.

sovereignty and windbaggery. And this difference must also, for the reasons given earlier, be connected in some way to the fact that the sovereign individual, unlike the feeble windbag, is, as I put it, socialized all the way down.

IV

Let's take the 'mastery' issue first. The suggestion that I want to make under this heading is essentially very simple; and it is this—that the sovereign individual is stirringly bloody-minded when it comes to standing 'security *for his own future*'. He refuses—and I think that this is the right way of putting the point—to treat other things as *not* being equal, even when, as a matter of fact, they aren't equal, or turn out not to be. So if ordinary promising requires intending, which in turn requires that the relevant intention is, *ceteris paribus*, executed, sovereign promising offers, as it were, to indemnify the promise against the *ceteris paribus* clause altogether, so that the intention is executed *regardless* (although there may be exceptions where promises come into conflict, either with one another or with other commitments: I return to this issue in Section VI, below). It is in this sense that the construal offered a moment ago wasn't completely wide of the mark: the sovereign individual does indeed promise only when he is sure that other things will be equal—not, as that construal had it, because he is confident that events will pan out, independently of him, in a manner conducive to the promise's being kept, but because he is committed to *take* them as equal—to treat them as equal—whether or not they actually are. He knows himself, as Nietzsche puts it, to be 'strong enough to maintain [his promise] in the face of accidents, even "in the face of fate" ' (GM II 2).[4]

An obvious example of this sort of thing, I take it, would be getting married. The betrothed make some really quite high-octane promises to one another, and they undertake to see these through, not just for as long as it is reasonably convenient to do so, but regardless—for better or for worse, in sickness and in health, for richer or for poorer, and so on. They indemnify their promises against the *ceteris paribus* clause. As Nietzsche himself puts it: 'The rationality of marriage lay in its principled indissolubility, which gave it an accent that knew *how to be heard* above the accidents of feeling, passion, and the distractions of the moment' (TI, Skirmishes, 39)—so that a marriage might be maintained 'even "in the face of fate" ' (for instance, in the face of temptresses, boredom, hair-loss, cellulite, football, children, fundamental incompatibility, menopause, snoring). And this is why the sovereign individual promises 'reluctantly, rarely, slowly'—even if, as in this case, 'rarely' might be a touch on the frequent side.

[4] For a somewhat different—and as I now think better—way to understand the sovereign individual here, see David Owen, this volume.

Sovereign promising, then, amounts to whole-hearted commitment to someone or something, to executing the relevant intention come what may.[5] And feeble windbaggery consists in frailty in the face, precisely, of 'the accidents of feeling, passion, and the distractions of the moment', and hence in a lack of mastery over 'circumstances', 'nature', and 'fate'. And in the case of marriage, or so Nietzsche plausibly suggests, such a lack is now overwhelmingly the norm:

> People live for today, people live very fast,—people live very irresponsibly: and this is precisely what people call 'freedom.' The things that *make* an institution into an institution are despised, hated, rejected: people think they are in danger of a new sort of slavery . . . : Witness *modern marriage*. It is clear that modern marriage is completely irrational: but this is an objection to modernity, not to marriage . . . . Modern marriage has *lost* its meaning,—consequently, it is being abolished. (TI, Skirmishes, 39)

Moreover, it is being abolished in the name of the windbag's version of freedom (as '*laisser aller*', letting go). 'But that', says Nietzsche, 'is a symptom of *decadence*: our modern concept of "freedom" [i.e. "laisser aller"] is one more proof of the degeneration of the instincts' (TI, Skirmishes, 41); which means that the 'freedom' of which the sovereign individual is conscious, even if it has 'penetrated to the profoundest depths and become instinct', can have nothing to do with letting himself go, with doing whatever he feels like. Rather, *his* freedom must be realized in—indeed be expressed in—his fidelity to his commitments, in seeing them through no matter what.

V

I said in Section II that what one does is constitutive of what one believes, desires, intends and so on. The point then was to steer off the thought that intending and acting might be radically separable from one another, so that, for instance, the former might be said to cause the latter. But the point can be extended. For it is not true, merely, that one can find out that one's sincere statement of intent was mistaken—when, for instance, one fails to act upon it although one could have done, and where other things were equal. It is also true that, in acting upon it, one can discover what the real or full nature of one's intention is. For example, I might say beforehand, and in all good faith, that my intention in getting the kids to bed on time is to make sure that they get enough sleep, but then, in the event, find that the *manner* in which I get them there is also (and not by accident) meant to reassert my authority over them, perhaps after a bad day. So I can not only discover that I didn't intend what I thought I did, but

[5] Again, the promises of a literal sovereign may well have just this character—in the case of a treaty, say, when the circumstances under which the sovereign undertakes to act (and undertakes perhaps even on behalf of his successors) may be, in practice, unforeseeable.

that I intended what I thought I didn't, or didn't know that I did. And this last point—that I can find out in acting exactly what I *do* intend—is, in a special set of cases, crucial to understanding the socialized dimension of the sovereign individual's promise-making.

We need to distinguish between three sorts of way in which one's actions can reveal the positive character of one's intentions (as opposed to revealing their absence). The first is merely confirmatory. Suppose that I set out to make a cup of instant coffee; I put some powder in the mug, boil the kettle, and add the contents of the latter to the former. Clearly I intended to make a cup of coffee, and did so. More impressive confirmation of my resolve might have been provided if, say, having discovered that the kettle had blown up or been stolen, I'd switched to the more time-consuming recourse of a saucepan. But in either case I end up with a made-by-me cup of coffee, and that is enough. (It is also enough if, upon switching to the saucepan, I decide to make soup instead. That I put some coffee-powder in a mug and show clear signs of wanting to boil something bears out the thought that I did indeed start out with the intention of making coffee; and that I end up with a bowl of soup shows that other things weren't equal and that I came to intend otherwise.)

The second way is exemplified in the putting-the-kids-to-bed case. My actions, we may suppose, confirm that I did indeed intend to get them to sleep; but they also reveal a previously unrecognized and perhaps subsidiary intention—namely, to remind everybody who's boss. Here I intended more than I thought, and what I do shows me to myself, quite possibly in an unnerving light (anyone who doubts this should simply adjust the example). This second way shares with the first an important feature: the actions involved, and the intentions of which they are the executions, are—or certainly can be—entirely *basic*, in the sense that one can non-trivially specify in advance what will count as success. I will have successfully executed my intentions if I end up with a made-by-me cup of coffee, punctually put-to-bed-by-me offspring, and a general sense that my authority has been reasserted (by me). In cases of this sort, where success-conditions are specifiable independently of any *particular* way of meeting them (it doesn't matter whether I use the kettle or a saucepan, for example, or whether I am peremptory in carrying them or in ordering them to bed), I shall say that those conditions are 'external' to the execution of the relevant intention.

The third way in which one's actions can reveal the positive character of one's intentions is different in this respect. Here, success-conditions are *internal* to the execution: what counts as success cannot be non-trivially specified in advance and independently of some particular way of meeting them. Nietzsche's own examples of cases of this sort are usually drawn from art, as in the following:

> Every artist knows how far from any feeling of letting himself go his most 'natural' state is—the free ordering, placing, disposing, giving form in the moment of 'inspiration'—and

how strictly and subtly he obeys thousandfold laws precisely then, laws that precisely on account of their hardness and determination defy all formulation through concepts. (BGE 188)

For a law to 'defy all formulation through concepts' is, precisely, for what counts as compliance with it to be unspecifiable. Of course, I might say independently and in advance that I intend to produce a stirring sonata or a memorable sonnet. But that is just to say that I intend to produce a successful one, which is trivial. In cases of this sort, then, non-trivial success-conditions emerge only during the course of the execution, so that the full character of my intention is revealed only in my (successfully) acting on it. (This is a point of overlap between this third way and the second: both, as I put it, can show me to myself.) In this sort of instance, therefore, I don't merely confirm that I intended to such-and-such, or find out that I intended more than I thought. Rather, I discover what it *was* to have intended to such-and-such, and so discover something about myself.

It is striking and uncoincidental that Nietzsche should connect cases of this kind to freedom, on the one hand, and to 'law' on the other. Indeed, he is insistent on the connection: the 'fact is', he says, 'that all there is or has been on earth of freedom, subtlety, boldness, dance, and masterly sureness, whether in thought itself . . . , or in rhetoric and persuasion, in the arts just as in ethics, has developed only owing to the "tyranny of . . . capricious laws" '—where a 'capricious' law is one that is socially constituted, one that is, in other words, the mark of an *institution*. '[A]nd in all seriousness', he continues, 'the probability is by no means small that this is "nature" and "natural"—and *not* that *laisser aller*' (BGE 188), which latter, as we have seen, is the 'modern concept of freedom'. Nietzsche's view, then, is that in activities of this kind—where one 'obeys thousandfold laws' that 'defy all formulation through concepts'—one is, first, properly to be called 'free', and, second, to be called free *in virtue* of one's internalization of the norms constitutive of the institution through which one acts. Freedom of this sort, therefore, is available *only* to someone who is, in the relevant respects, socialized, as I put it, all the way down.[6]

VI

If what I have argued in the previous two sections is plausible, we now have a pretty good sketch of what the sovereign individual's promise-making involves. He has 'mastery' over 'circumstances', 'nature', and 'fate' because he is ready to indemnify his promise against the *ceteris paribus* clause. And the most

[6] This and the previous paragraph touch on themes that I have explored more thoroughly elsewhere, in 'Nietzsche on Art and Freedom' (2007). *European Journal of Philosophy* 15/2: 204–24, and in my Introduction to Friedrich Nietzsche, *The Anti-Christ, Ecce Homo, Twilight of the Idols* and other writings (Cambridge: Cambridge University Press, 2005).

characteristic *content* of his promise is an intention whose success-conditions are internal to the execution—that cannot, that is, be non-trivially specified in advance and independently of some particular way of meeting them.

With this in mind, let's return to an earlier example, and imagine someone promising, in good faith, to be a decent husband. It is true that there are some independently specifiable success-conditions here (although they are defeasible). Respect is presumably necessary, for example, as are caring for the other person's interests and not betraying them, say. But exactly what might *count* as betrayal, or what caring for the other person's interests might *look* like in this case—or even whether *these* things are what is at issue—cannot be specified independently of the particular marriage that it is, of the circumstances, history and personalities peculiar to it, and of how those things unfold or develop over time. It is, in other words, perfectly possible that everything I do is, as it were, strictly speaking not disrespectful or inconsiderate or disloyal, and yet that I fail to be any good as a husband—I am true to the letter but miss the spirit, as we might say. And being true to the letter, we must suppose, is something that even a (relatively high-grade) windbag ought to be able to bring off, perhaps, indeed presumably, entirely unaware that the spirit has eluded him. (To say that there is a spirit here to be eluded by is one way of bringing out what is distinctive of fully sovereign promising. In more basic kinds of undertaking—coffee-making, for instance—there *is* no spirit; the letter is everything, and is indefeasible. I can gather what is required *merely* by being given a set of instructions; and I may, if I'm a bit mad, determine to see these through regardless, 'even "in the face of fate" '. But no such instructions are available in the case of fully sovereign promise-making: the internal character of the relevant success-conditions, together with the nature of the sovereign individual's freedom, precludes this.)

If this is right, then good faith in this context has to do with discovering, fine-tuning, recognizing what the spirit of the relevant undertaking in fact turns out to require, all of which is a matter of what Aristotle would have called 'perception'. And maybe, in some cases, good faith will eventually require that the character of the undertaking be revised, perhaps as a result of finding that the original one was unrealizable, given the circumstances as they turned out to be—of finding, in fact, that there was not, after all, an intention of *that* shape to be executed. And here we might agree to be just good friends, say, in which case many of the independently specifiable success-conditions relevant to being a good husband may well stay in place, but a different conception of what it might be to satisfy them emerges as the ends of marriage give way to those of friendship.

But perhaps the sovereign individual would never find himself in this latter position. It suggests, after all, either an insufficient mastery over 'fate' or else a promise entered into too impetuously. But it seems to me to be an alternative worth keeping on the table even so, for two reasons. First, it ensures that the sovereign individual represents a genuinely *human* possibility, rather than a merely super-human one: human infallibility is an idle fantasy, after all. And

second, it ensures that the sovereign individual is continuous, in at least one important respect, with the figure of the artist, upon whom Nietzsche draws in his characterization of properly free agency. When Brahms, for example, found that what he'd thought would be a string quintet had to be a piano quintet instead—that the musical facts as they turned out to be required that—he was not being any kind of windbag.[7] Rather he was remaining true to his wider commitment to produce decent music; and good faith, in this instance, spoke against insisting that what could never be a worthwhile work for strings alone must nevertheless be *made* to be one. And the sovereign individual likewise: his wider commitment, or so we might imagine, to worthwhile human relationships obliges him to recognize that this one, at any rate, *cannot* be a marriage, despite what he had in all good faith thought. There are different levels here, and anything worth calling free agency must be allowed to range across them, if we are not to be left with the merest fanaticism.

I should say a little more about this last point. Except for just now, I have tended to speak as if the sorts of commitment that are characteristic of the sovereign individual were, first, mutually compatible and, second, all of equal significance. But it is clear that none of this need be the case. In the examples just given, for instance, narrower commitments—to write a string quintet, to make a go of a marriage—were trumped, in effect, by higher-order commitments—to decent music, to decent relationships—of which the narrower ones were special cases. So we should expect there to be the occasional conflict (however 'reluctantly, rarely, slowly' the sovereign individual promises); and we should expect these to be resolved in accordance with some sort of hierarchy among the commitments involved.[8] Indeed, to 'maintain' an '*order of rank* in the world, among things themselves', Nietzsche tells us, is precisely what that 'high spirituality', which 'exists only as the ultimate product of moral qualities', of the 'synthesis of all those states which are attributed to the "merely moral" man', achieves (BGE 219). And this—the maintenance of an '*order of rank*'—is something else that the modern age isn't much good at, or so Nietzsche suggests:

> Nothing seems rarer to me today than genuine hypocrisy . . . . Hypocrisy belongs to an age of strong belief: where people do not give up their belief even when they are *forced* into pretending to adopt another. People will just give up their belief these days; or, more likely, they will take on an additional one,—in either case they stay *honest*. No doubt many more convictions are possible these days than in the past: possible means allowed,

[7] Actually, we have to take Brahms's word for this: he destroyed the string version. (And before the work eventually became the one that we know [Op.34], it had a surviving—and later published—incarnation as a work for two pianos, although I personally can't see why Brahms didn't destroy that one too.)

[8] Conflicts may arise between undertakings with the same general telos—as, for instance, in the music and relationships examples—or with different ones—as when, for instance, my unshakable commitment to meet you for a drink is stymied by the need to rescue someone from mortal danger. In either kind of case, though, what matters is that one undertaking is (or can be) recognized as taking precedence over the other.

it means *harmless*. This is where tolerance of yourself comes from.—Tolerating yourself allows for multiple convictions: and these live together in harmony,—like everything else, they are careful not to compromise themselves. How do you compromise yourself these days? By being consistent . . . . By being genuine . . . I greatly fear that modern man is just too indolent for certain vices: so that they are actually dying out. (TI, Skirmishes, 18)

The 'genuine' hypocrite, then, is at least properly clear about and committed to his higher-order convictions (i.e. he *has* a properly formulated '*order of rank*') and that, *pro tanto*, puts him on the side of the angels—that is, on the side of the sovereign individual, one who has his own 'standard of value'.

These points prompt a refinement of the account of sovereign promising developed so far. The sovereign individual, I have said, indemnifies his promises against the *ceteris paribus* clause. We can now see, though, that this can only be strictly speaking true of promises at the top end of his order of rank. Lower-ranking promises, although he might indemnify them against every other inequality, cannot be indemnified in the case of conflict with promises higher up the scale. Where there *is* such a conflict, the lower-ranking promise must be broken—*not*, as in the windbag's case, because of 'the accidents of feeling, passion, and the distractions of the moment', but because of the sovereign individual's genuine (i.e. sovereign) commitment to a particular hierarchy of values.[9] And with this qualification in place, we can say that his '*right*' to make promises consists in his capacity to commit himself whole-heartedly to undertakings whose character is inconceivable except in the context of the social world, and of the formulable *and* unformulable laws comprising it, from which they draw their sense. Whereas the windbag, by contrast, is either fickle, inadequately socialized, or both.

VII

I said at the beginning that I'd return to the assumption that the sovereign individual stands outside of traditional morality, and it is time to make good on that. What I particularly want to do is to spell out, in light of the foregoing, what Nietzsche's remark that the sovereign individual is 'autonomous and supramoral (for "autonomous" and "moral" are mutually exclusive)' amounts to. That the remark is meant to be provocative goes without saying: Kant, after all, had identified the autonomous with the moral (and both, for good measure, with freedom). So the sovereign individual's position is, at the very least, supposed

[9] Nietzsche is not inclined to be prescriptive about the content of such a hierarchy. As he puts it in a related context, what matters is that 'the constraint of a single taste' should have 'governed and formed everything large and small. Whether this taste was good or bad is less important than one might suppose, if only it was a single taste!' (GS 290).

to be incompatible with Kantian morality. The sovereign individual is also said to be 'liberated again from the morality of custom', so he occupies a position outside of that as well.

Let's take the latter point first. To be part of the 'morality of custom', recall, consists in 'finally' remembering 'five or six "I will not"s', as a condition of enjoying 'the advantages of society'. It is, in other words, a matter of accepting, and acting upon, a set of instructions, mere compliance with which is sufficient. And this, as we saw a moment ago, is to fall short of what the sovereign individual undertakes to do: he complies with laws that defy 'formulation through concepts', and so which cannot be exhaustively translated into any set of instructions. So translated, the relevant instructions might, perhaps, capture the letter of some of his laws; but they cannot capture the spirit. It is quite easy, then, to see in what sense the sovereign individual is 'liberated' from 'the morality of custom'—i.e., is made, among other things, freer than that—since his way of acting cannot be captured in *any* mere list of 'I will not's', however long.

And the situation with Kantian morality is essentially the same. In *The Gay Science* Nietzsche makes the following remark: 'Anyone who still judges "in this case everyone would have to act like this" has not yet taken five steps toward self-knowledge' (GS 335), a remark clearly directed at Kant, and meant to ridicule the categorical—that is, the universal—character of his imperatives. If I find that the maxim of my action cannot be universalized without contradiction, I have identified an absolute prohibition, an unconditional 'I will not'. I have, in other words, stopped short at a formulable instruction that might be fully obeyed by anyone, the feeble windbag included. The spirit—for all of Kant's talk of the good will's identifying itself with the deliverances of reason—has gone missing without trace. It is as if Brahms, having completed his Paganini variations, had declared that *anyone*, confronted with this theme and that brief, would have had to go about it in the same way. But no: Brahms (perhaps) had to; yet Rachmaninov certainly didn't. And for Brahms to have made such a declaration would have been for him to mistake a claim that applied, at most, to himself for something quite different[10]—which is why the Kantian moralist 'has not yet taken five steps toward self-knowledge'.

To be autonomous, then, in Nietzsche's sense, is to have the capacity to commit oneself to acting on intentions whose success-conditions are (at least partly) internal to the execution, and hence in coming to discover the determinate nature of those intentions (i.e. to acquire self-knowledge). It is, in other words, to do altogether *more* than simply comply with a set of prescriptions. And since, in Nietzsche's view, 'morality'—understood as some set or other of prescriptions, mere compliance with which is sufficient—cannot make sense

[10] Cf. Nietzsche's remark in *Ecce Homo* that he had to free himself of the (inherited) conviction that he was just 'like everyone else' ('Why I am So Clever', 2).

of or room for such a capacity, ' "autonomous" and "moral" ' must, indeed, be 'mutually exclusive'. And this is so irrespective of whether the compliance required by 'morality' is merely outward, as in the case of the 'morality of custom', or is inward too, as in the case of Kantian morality: for no amount of fervent inward identification with the moral law will bring it about that that law suddenly becomes unformulable, or that the success-conditions of one's intention to comply with it suddenly become internal to the execution. Indeed, Kantian 'autonomy', as Nietzsche construes it, is merely a technical name for a very particular sort of self-misunderstanding—the sort that arises, precisely, when one judges that ' "in this case everyone would have to act like this" '.[11]

The Nietzschean notion of autonomy, by contrast, and the conception of freedom that goes with it, has to do with *coming* to understand oneself. One is not merely constrained by an externally specifiable list of prescriptions, of 'I will not's—although one is constrained by these. One is also, more importantly, *not* constrained by the thought that mere compliance with such prescriptions is all that there is or could reasonably *be*. Instead, one treats the domain traced out by these prescriptions as the space within which an exploration of oneself, and of the socialized world that has, in the limiting case, become 'natural' and even instinctive, is to be conducted. For only thus are the great undertakings—artistic, philosophical, political, personal—for which Nietzsche reserves his respect so much as conceivable. And it is as a marker of the real possibility of these, if what I have argued here is correct, that he writes of the sovereign individual, and of his '*right*' to make promises.[12]

[11] The passage (GS 335) goes on, in a way that many have been misled by, to insist that 'there neither are nor can be actions that are the same; that every action that has ever been done was done in an altogether unique and irretrievable way . . . ; [that] every action is unknowable'. It can seem—and in all honesty it sometimes seemed to Nietzsche (see, e.g., GS 354)—as if the interesting point here were that, because no two actions are exactly alike (which is trivially true at most), we have a real problem with respect to our knowledge of them, and hence have a real problem with our understanding of ourselves. But the deeper point is more or less the opposite of this: it is that, if we concentrate only on what distinct actions have in common, we end up in the realm of the merely formulable; and if we confine ourselves to *that*, we do indeed misunderstand ourselves—for example, we fail to see that what *this* undertaking requires of *us*, right now, is acting in this way rather than that (although both ways have it in common, as it might be, that they are compliant with the injunction: do not criticize her clothes; or: do not get out of E minor too quickly). This is why Nietzsche takes the following to be relevant to his point: 'all regulations about actions relate only to their coarse exterior (even the most inward and subtle regulations of all moralities so far)' (ibid.)—where 'regulations', clearly enough, equal 'prescriptions'; and where the epithets 'inward' and 'subtle' plausibly apply to Kantian morality. (It is a constant wonder that Nietzsche has nothing of worth to say about Aristotle.)

[12] This paper began life as two separate papers—one delivered at the 2004 conference of the Friedrich Nietzsche Society at the University of Sussex, the other as a response to an essay by Robert Pippin at a conference, 'Robert Pippin: Modernism and the Ambitions of Philosophy', held at Southampton in 2005. I am grateful to participants in both events for their comments. I am also grateful to Maria Alvarez and David Owen, both of whom were indispensable to the process of turning two quite messy papers into this single one, which is, I hope, a good bit tidier than either of the originals.

REFERENCES

WRITINGS BY NIETZSCHE

On the Genealogy of Morals, in *On the Genealogy of Morals and Ecce Homo*, trans. Walter Kaufmann and R. J. Hollingdale (New York: Vintage, 1969).

Beyond Good and Evil, trans Walter Kaufmannn (New York: Vintage, 1966).

The Gay Science, trans. Walter Kaufmann (New York: Vintage, 1974).

Twilight of the Idols, in *The Anti-Christ, Ecce Homo, Twilight of the Idols and other writings*, ed. Aaron Ridley, trans. Judith Norman (Cambridge, Cambridge University Press, 2005).

OTHER WORKS CITED

Havas, Randall (2000). 'Nietzsche's Idealism', *Journal of Nietzsche Studies* 20, 90–9.

Pippin, Robert (2004). 'Lightning and Flash, Agent and Deed (GM I:6–17)', in O. Höffe, (ed.), *Friedrich Nietzsche, Genealogie der Moral* (Berlin: Akademie Verlag), 47–63.

Ridley, Aaron (2005). 'Introduction', in Aaron Ridley and Judith Norman (eds.), *The Anti-Christ, Ecce Homo, Twilight of the Idols and other writings* (Cambridge: Cambridge University Press).

——(2007). 'Nietzsche on Art and Freedom', *European Journal of Philosophy* 15:2, 204–24.

This page intentionally left blank

10

Autonomy, Self-Respect, and Self-Love: Nietzsche on Ethical Agency[1]

David Owen

> *It is the music in our conscience, the dance in our spirit that makes all the puritanical litanies, all the philistinism and moral sermons sound so dissonant.*
>
> (BGE 216)

It has been perceptively argued that 'Kant tried to resolve the tension between law and love that runs through early modern moral philosophy' by 'distinguishing morality into a domain of right and a domain of virtue'.[2] Yet while Kant may be said to have succeeded in reducing the tension between law and love in modern moral philosophy, it would be overstating the case to see him as achieving a resolution of this tension. Kant's effort to accommodate law and love through a structural division between the realms of right and of virtue allows him to stress the moral significance of both self-respect and respect for others, and to make room for the love of others that develops from the (wide) duty to promote their happiness, yet he adopts a highly critical stance towards self-love throughout his moral philosophy. Self-love—conceived in a distinctly modern sense as the stance of *psychological egoism* expressed in acting on maxims oriented to the pursuit of one's own happiness—is opposed to *morality* as a threat, a challenge, a danger; an opposition whose stakes are intensified by Kant's acknowledgement that self-love is an ineliminable feature of finite desiring (hence happiness-seeking) creatures such as human beings and, consequently, that one can never be sure

[1] An early version of this chapter was presented to a symposium organized by Ken Gemes at Birkbeck in late 2006 and I am grateful to those present for their comments. A later version was presented to the 'Nietzsche in New York' meeting in March 2007 and I benefited from comments by Christa Davis Acampora, Daniel Conway, Robert Guay, Lawrence Hatab, and Tracy Strong. Still later versions received the constructive criticism of Russell Bentley, Peter Niesen, Simon Robertson, and Aaron Ridley. I am particularly grateful to Simon for a very extensive set of comments that I have not been able to do full justice to in my revisions and to Aaron for providing some of the necessary foundations for this argument. I would also like to thank Ken Gemes and Simon May for helpful editorial advice.

[2] Schneewind (1998: 528).

that actions complying with one's moral duties are solely motivated by respect for the moral law.

This modern distinction between morality and egoism expressed by Kant in terms of the opposition of the moral law and self-love is mischievously exploited by Nietzsche for its rhetorical shock value in *Daybreak* when he presents his project of re-evaluation as, in part, oriented to the following task: 'we shall restore to men their goodwill towards the actions decried as egoistic and restore to these actions their *value—we shall deprive them of their bad conscience*!' (D 148). A slightly contrasting and rather more diagnostic stance on this topic is adopted in the opening essay of *On the Genealogy of Morality* in which Nietzsche argues that it is only with the decline of aristocratic value-judgments that 'this whole antithesis between "egoistic" and "unegoistic" forces itself more and more on man's conscience' and 'even then' he remarks:

> it takes long enough for this instinct to become sufficiently dominant for the valuation of moral values to become enmeshed and embedded in the antithesis (as is the case with contemporary Europe, for example: the prejudice which takes 'moral', 'unegoistic' and '*desinteresse*' as equivalent terms already rules with the power of a 'fixed idea' and mental illness). (GM I 2)

What is clear from both of these remarks is that Nietzsche has an objection, whose nature we will be concerned to determine, to the modern framing of ethics in terms of the opposition of morality and egoism—and at least part of his re-evaluation of moral values is to involve recasting ethics in a way that avoids this opposition by reopening the debate on law and love that Kant had sought to conclude. To situate this discussion, however, it will be worthwhile to begin with some brief reflections on Aristotle and Kant in relationship to the topics of self-love and self-respect.

I

In Book IX of the *Nichomachean Ethics*, Aristotle offers two discussions of self-love, the first of which aligns self-love with his discussion of virtue (or character) friendship and the second of which distinguishes 'true' self-love from 'pseudo' self-love. For the purposes of this chapter, it will be appropriate to begin with the second of these before turning to the first.

Acknowledging that the description of someone as a lover of self is popularly taken as a mark of criticism, Aristotle stresses that those 'who use the term as one of reproach ascribe self-love to people who assign themselves the greater share of wealth, honours, and bodily pleasures; for these are what most people desire, and busy themselves about as though they were the best of all things' and notes that 'those who are grasping with regard to these things gratify their

appetites and in general their feelings and the irrational element of the soul' (NE IX.8). In this regard, Aristotle endorses the critical stance of the popular view. However, his argument proceeds to distinguish such 'apparent' or 'pseudo' self-love from the *true* self-love which is manifest when a man 'assigns to himself the things that are noblest and best, and gratifies the most authoritative element in himself and in all things obeys this' (NE IX.8). Thus, in the context of Aristotle's broader account of *eudaemonia, true* self-love is to be seen as an acquired relation to self that consists in appropriating to oneself (committing oneself to the service of) what is noblest by submitting to the rule of reason and, hence, identifying one's well-being with living a virtuous life. *True* self-love, thus, consists in the disposition to love what is noble; the contrasting state to such *true* self-love is the *pseudo* self-love exhibited in the self-interested pursuit of pleasure, that is, the identification of happiness with the satisfaction of one's (irrational) appetites. In sum: *true* self-love is the natural disposition of the virtuous man in the sense both that acting virtuously promotes the disposition of *true* self-love (since acting virtuously is acting on reasons which express the virtues, that is, what is noblest and best) and that the disposition of *true* self-love promotes virtuous action (since it is the disposition to act on reasons which express the virtues).

Aristotle's other discussion of self-love advances the argument that self-love is the basis of friendship since those features that define friendship such as valuing one's friend for his own sake, identifying with his interests and well-being and, hence, grieving and rejoicing with him are also true of the good man's relationship to himself:

> For his opinions are harmonious, and he desires the same things with all his soul; and therefore he wishes for himself what is good and what seems so, and does it (for it is characteristic of the good man to work out the good), and does so for his own sake (for he does it for the sake of the intellectual element in him, which is thought to be the man himself); and he wishes himself to live and be preserved, and especially the element by virtue of what he thinks. (NE IX.4)

By contrast, 'the bad man does not seem to be amicably disposed even to himself, because there is nothing in him to love' (NE IX.4). From the context, it seems clear that Aristotle is focusing on what he will shortly specify as *true* self-love and on that (highest) species of friendship, namely, virtue or character friendship. The three central claims for our purposes are thus (a) that *true* self-love is the basis of character friendship, (b) that *true* self-love is structurally analogous to character-friendship and (c) that this structural analogy is constituted by the presence in both self-love and character-friendship of a reason-governed concern for the well-being of the self or other for its own sake. It is no doubt the case that for modern sensibilities, as Nietzsche acutely diagnosed, to be concerned for one's own well-being for one's own sake sounds suspiciously egoistic, however,

be that as it may, the helpfulness of Aristotle's analogy is that it clarifies certain features of *true* self-love by directing our attention to the nature of the self's relationship to itself exhibited in *true* self-love as one in which one acts as a character-friend to oneself. The salience of this analogy is given by considering two points. First, that to stand as a character-friend to another is to value the other in virtue of their disposition to love and pursue what is noble. This helps to explain why Aristotle sees self-love as the basis of character-friendship in that character-friendship is characterized by valuing the disposition to love what is noble in another and, hence, is predicated on valuing the disposition to love what is noble, and valuing the disposition to love what is noble is to manifest *true* self-love. Second, a character-friend is in a privileged position with respect to their friend in that they are entitled—indeed, when appropriate, compelled—to act as a critical interlocutor, that is, a *conscience*, to their friend when the friend threatens to stray, or does depart, from conduct that is noble. Other things being equal, the friend who overlooks your lapses and occasional viciousness is not acting as a character-friend. In this respect, we might distinguish between the *true* and *pseudo* character-friendship in a way that parallels the distinction between *true* and *pseudo* self-love. The former, precisely in being concerned for your well-being for its own sake, will be prepared to act as a conscience for you in a way that the indulgence extended by the latter signally fails to do. To exhibit *true* self-love thus necessarily requires that one stand in a critical relationship to oneself, make demands of oneself with respect to the character of one's conduct and hold oneself to account in respect of one's success or failure to meet such demands, not least in terms of appropriately experiencing the emotions of pride or shame that attend one's successes and one's failures. By contrast, *pseudo* self-love is a state of indulgent pleasure-seeking that leads, fatefully, to a condition in which considerations of ethical evaluation have no grip on the agent at all, the condition of shamelessness.

In these brief comments on Aristotle, I have stressed the distinction between *true* and *pseudo* varieties of self-love not least because, despite what I take to be the intuitive plausibility of such a distinction, Kant will deny it any place in his moral philosophy. His stance is stated succinctly in the section on 'Egoism' in *Anthropology from a pragmatic point of view*:

> . . . the moral egoist limits all purposes to himself; as a eudaemonist, he concentrates the highest motives of his will merely on profit and his own happiness, but not on the concept of duty. Because every other person has a different concept of what he counts as happiness, it is exactly egoism which causes him to have no touchstone of a genuine concept of duty which truly must be a universally valid principle. All eudaemonists are consequently egoists. (AP Bk I 2)

Kant spells out the reasons for adopting this stance in his *Critique of Practical Reason* by, first, specifying his view of the character of self-love and, second, its relationship to the moral law. Thus:

All the inclinations together (which can be brought into a tolerable system and the satisfaction of which is then called one's own happiness) constitute regard for oneself (*solipsismus*). This is either the self-regard of *love for oneself*, a predominant *benevolence* toward oneself (*Philautia*) or that of *satisfaction with oneself* (*Arrogantia*). The former is called, in particular, *self-love*; the latter, *self-conceit*. (CPR 5:73)

The relationship between *self-love* and *self-conceit*, as well as their relationship to the moral law, is then explained thus:

This propensity to make oneself as having subjective determining grounds of choice into the objective determining grounds of the will in general can be called *self-love*; and if *self-love* makes itself lawgiving and the unconditional practical principle, it can be called *self-conceit*. Now the moral law, which alone is truly objective (namely, objective in every respect), excludes altogether the influence of self-love on the supreme practical principle and infringes without end upon self-conceit, which prescribes as laws the subjective conditions of self-love. (CPR 5: 74)

Thus, Kant argues that, in contrast to the moral law, self-love cannot provide an objective basis for morality. It might be objected that at this stage that Kant simply identifies self-love *per se* with what, for Aristotle, would count as *pseudo* self-love—and there is no doubt that Kant does tend to see self-love in this way. However, the underlying reason for Kant's rejection of the distinction between *true* and *pseudo* self-love as a way of distinguishing the moral and non-moral agent is not that he denies that there can be individuals with a settled and effective disposition to, say, beneficence. It is, rather, that he denies that actions solely based on the inclination to beneficence have moral worth as the following example from the *Groundwork* makes clear:

To be beneficent where one can is a duty, and besides there are many souls so sympathetically attuned that, without any other motive of vanity or self-interest they find an inner satisfaction in spreading joy around them and can take delight in the satisfaction of others so far as it is their own work. But I assert that in such a case an action of this kind, however it may conform with duty and however amiable it may be, has nevertheless no true moral worth but is on the same footing with other inclinations, for example, the inclination to honour, which, if it fortunately lights upon what is in fact in the common interest and in conformity with duty and hence honourable, deserves praise and encouragement but no esteem; for the maxim lacks moral content, namely that of doing such actions not from inclination but *from duty*. (G 4: 398)

Illuminatingly Kant continues:

Suppose, then, that the mind of this philanthropist were overclouded by his own grief, which extinguished all sympathy with the fate of others, and that while he still had the means to benefit others in distress their troubles did not move him because he had enough to do with his own; and suppose that now, when no longer incited to do it by any inclination, he nevertheless tears himself out of their deadly insensibility and does the action without any inclination, simply from duty; then the action first has its genuine moral worth. (4: 398)

These passages suggest a general point concerning the opposition of inclination (roughly, motives for actions that pertain to our sensuous bodily nature) and duty (roughly, motives for action that pertain to our rational nature) in Kant's moral psychology. It is this: acting according to inclination is both (a) dependent on the contingent circumstances that inform our propensity to be characterized by the presence or absence of particular inclinations and (b) orients us to the effects of our willing, to seeing our will as good in relation to 'its fitness to attain some proposed end', whereas acting according to duty identifies the goodness of a good will as relating to the activity of willing in and of itself, that is to the form of volition expressed in agency independent of either the contingent inclinations of the agent at the time of acting or the effects of such agency. Another way of putting this point, and understanding Kant's motivation for adopting it, is to see Kant as arguing that since acting according to inclination is both a product of luck (our current empirical constitution) and exposed to luck (not least in respect of both the action of events in the world on our empirical constitution and on the outcomes of our actions), it cannot be conceived of as an *objective* (in all the relevant senses) basis for the ascription of moral worth. By contrast, acting according to duty can be conceived as an *objective* basis for the ascription of moral worth and, hence, respect in that it is immune from luck since the will is determined by nothing other than 'the conformity of actions as such with universal law, which alone is to serve the will as its principle, that is, *I ought never to act except in such a way that I could also will that my maxim should become a universal law*' (G 4: 402). Respect for the moral law, exemplified in acting according to duty, is thus the basis of the disposition of self-respect (as well as that of respect for others) that characterizes the morally autonomous agent.[3]

These considerations concerning Kant's arguments on the foundation of morality should not be taken to imply that he rejects all forms of self-concern. On the contrary, Kant's stress on the (wide) duty to achieve one's own perfection indicates the centrality of an orientation to the cultivation of oneself as a moral agent within his moral philosophy (cf. MM 6: 393). However, it remains the case that in the contrast between Aristotle's account of the virtuous man characterized by the disposition of *true* self-love and Kant's account of the morally autonomous agent characterized by the disposition of self-respect, we can see vividly the tension between the perspectives of love and of law that inform the tradition of European moral philosophy. The thesis advanced in the remainder of this chapter is that Nietzsche's ethics is an attempt to dissolve this tension. The argument for this claim proceeds by focusing, first, on Nietzsche's account of autonomy and self-respect and, second, on the relationship of this account to his understanding of self-love.

[3] For subtle and sympathetic reflection on this topic in Kant's moral philosophy, see Hill Jr. (1991).

II

To begin reflecting on Nietzsche's account of ethical autonomy and self-respect, I'll focus on his account of the figure of the sovereign individual in the second essay of *On the Genealogy of Morals*. In certain respects, this is a controversial decision since it has recently been argued that the account of the sovereign individual does not represent Nietzsche's own account of ethical agency. Consequently, in setting out an analysis of Nietzsche's discussion of this figure, I will also respond to such criticisms and demonstrate how my analysis of Nietzsche's account of the sovereign individual is fully compatible with his later—and explicitly endorsed—account of freedom in *Twilight of the Idols*.

The second essay of the *Genealogy* begins with Nietzsche considering the conditions under which human beings become capable of making and holding to promises (read: commitments). His starting point is to note that the ability to make commitments presupposes a variety of capacities:

> To think in terms of causality, to see and anticipate from afar, to posit ends and means with certainty, to be able above all to reckon and calculate! For that to be the case, how much man himself must have become *calculable, regular, necessary*, even to his own mind, so that finally he would be able to vouch for himself *as future*, in the way that someone making a promise does! (GM II 1)

Thus the question becomes that of how the task of '*making* man to a certain extent necessary, uniform, an equal among equals, regular and consequently calculable' (GM II 2) is accomplished? Nietzsche's hypothesis is that this is achieved in the prehistory of humanity 'by means of the morality of custom and the social strait-jacket' (GM II 2). By this, Nietzsche means simply that it was through the morality of custom and the social strait-jacket that human beings became creatures about whom one could legitimately have normative expectations concerning their activity since they have acquired the capacity not only for second-order desires but also for second-order volitions. When Nietzsche speaks of the development of 'a real *memory of the will*' as 'an ongoing willing of what was once willed [. . .] so that between the original "I will", "I shall do", and the actual realization of the will, its *enactment*, a world of new and strange things, circumstances, even other acts of will may safely intervene, with causing this long chain of will to break' (GM II 1), he is stressing the centrality of second-order volitions to personhood. At the end of this process

> where society and its morality of custom finally reveal the *end* to which they were merely a means: there we find as the ripest fruit on their tree the *sovereign individual*, the individual who resembles no one but himself, who has once again broken away from the morality of custom, the autonomous supra-moral individual (since 'autonomous' and 'moral' are

mutually exclusive)—in short, the man with his own independent, enduring will, whose *prerogative it is to promise*. (GM II 2)[4]

In this initial introduction of the figure of the sovereign individual, Nietzsche presents an image of *maturity* ('the ripest fruit') articulated in terms of *individuality* ('the individual who resembles no one but himself'), where such individuality is linked to *autonomy*, that is, being able to impose binding norms on oneself ('the man with his own independent, enduring will, whose *prerogative it is to promise*'). This figure represents, it seems, the concept of the autonomous individual who is not bound by moral rules as customary constraints, but as the freely endorsed commitments through which he gives expression to his own character. However, to specify this image of *maturity* further and to address the issue of whether Nietzsche may be said to endorse it requires that we attend more closely to Nietzsche's further delineation of the figure of the sovereign individual as one who exhibits 'a proud consciousness, tense in every muscle, of *what* has been finally achieved here, of what has become incarnate in him—a special consciousness of power and freedom, a feeling of the ultimate completion of man' (GM II 2). Nietzsche continues:

> This liberated man who has the *prerogative* to promise, this master of *free* will, this sovereign—how should he not be aware of his superiority over everything which cannot promise and vouch for itself? How should he not be aware of how much trust, how much fear, how much respect he arouses—he '*deserves*' all three—and how much mastery over circumstances, over nature, and over all less reliable creatures with less enduring wills is necessarily given into his hands along with this self-mastery. (GM II 2)

Given this consciousness, Nietzsche concludes this passage by drawing attention to the criteria of evaluation of persons deployed by the sovereign individual:

> The 'free' man—the owner of an enduring indestructible will—possesses also in the property his *measure of value*: looking out at others from his own vantage point, he bestows respect or contempt. Necessarily, he respects those who are like himself—the strong and reliable (those with the *prerogative* to promise), that is, anyone who promises like a sovereign [. . .] who is sparing with his trust, who *confers distinction* when he trusts, who gives his word as something which can be relied on, because he knows himself strong enough to uphold it even against accidents, even 'against fate'. Even so, he will keep the toe of his boot poised for the cowering dogs who promise without that prerogative, and hold his stick at the ready for the liar who breaks his word the moment he utters it. The proud knowledge of this extraordinary privilege of *responsibility*, the consciousness

[4] I draw here on the forthcoming revised Cambridge translation by Carol Diethe. The relevant German phrase in the passage cited is 'das *versprechen darf*' which has been variously translated in terms of the *right* to promise (Kaufmann/Hollingdale), being *entitled* to promise (Smith) and being *permitted* to promise (Clark/Swenson) with the last of these being perhaps the most literal translation but Diethe's translation now strikes me as the best way of emphasizing that this is a *power* without inclining one to notions of right. I am grateful to Christa Davis Acampora for pressing me on this point, although we do not agree on its implications for the interpretation of the salience of the figure of the *sovereign individual* in Nietzsche's argument; for her view, see Acampora (2006).

of this rare freedom, this power over oneself and over fate has sunk down into his innermost depths and has become an instinct, a dominant instinct—what will he call it, this dominant instinct, assuming he needs a name for it? About that there can be no doubt: this sovereign man calls it his *conscience* . . . . (GM II 2)

One initial concern raised by Hatab[5] and echoed by Acampora[6] in relationship to the argument that the sovereign individual represents Nietzsche's account of ethical maturity is that Nietzsche seems to align the sovereign individual with 'free will'—and yet we know that Nietzsche was himself highly critical of the doctrine of freedom of will 'in the superlative metaphysical sense' on the grounds that it requires appeal to the idea of the agent as *causa sui*. Moreover, as Hatab and Acampora both argue, this alignment of the sovereign individual with free will seems to imply a commitment to *moral* accountability, responsibility and retribution that Nietzsche can be said to have opposed.[7] Now it is certainly true that if it were the case that the sovereign individual must be construed in terms of 'freedom of the will' and '*moral* responsibility' in the superlative metaphysical sense that Nietzsche dismisses as fantasy (BGE 21), this would undermine the claim that the sovereign individual represents Nietzsche's own account of ethical maturity. However, this conclusion only follows if one fails to distinguish between two analytically distinct ways of reflecting on the topic of free will which Ken Gemes has termed '*deserts* free will' and '*agency* free will'.[8] *Deserts* free will is oriented to the issue of justified blame and is represented in the form of debates over free will and determinism. *Agency* free will is oriented to the question of what distinguishes mere doings from actions and is represented in the form of debates over the conditions which must be met for one to be an agent, for one's doings to be deeds. Nietzsche's rejection of *deserts* free will (and indeed of the whole free will/determinism debate) does not imply a rejection of concern with the topic of *agency* free will; on the contrary, Nietzsche exhibits a continuous concern with the question of the conditions under which doings (events) are deeds (action), a central element of his critique of the slave revolt in morality being that it articulates the conceptual resources for, and motivates, a systematically distorted understanding of ethical agency.[9]

Of course, once it is recognized that Nietzsche's discussion of the sovereign individual is a discussion of *agency* free will and, hence, that his description of the sovereign individual as a 'master of *free* will' entails that this discussion is offering an account of the conditions of ethical agency, it can also be seen that Nietzsche is indeed endorsing an account of responsibility that is tied to autonomy. But isn't Nietzsche, as Hatab insists,[10] highly critical of the Kantian ideal of moral autonomy and its correspondingly moralized account of responsibility? This is

[5] Hatab (1995: 37–8). [6] Acampora (2006).
[7] Hatab (1995: 38) and Acampora (2006). [8] Gemes (2009). See also Guay (2002).
[9] See Pippin (2004) and Owen (2007 ch. 5) for apposite remarks on this issue.
[10] Hatab (1995: 37–8).

certainly true. Yet it does not follow from rejection of the Kantian ideal of moral autonomy and the blame-focused account of responsibility that is tied to it that Nietzsche also rejects the idea of ethical autonomy exemplified by the figure of the sovereign individual; on the contrary, everything here hangs on the nature of the condition that the sovereign individual exemplifies—the topic to which we now turn.

In one respect, the evaluative contrast drawn in Nietzsche's discussion of the sovereign individual is between those who are entitled to *represent* themselves 'to others as holding certain beliefs or attitudes' or commitments and those who 'do not have the same right to speak in this way on their own behalf'.[11] As Lovibond puts it:

> Only on condition that I have, for example, sufficient self-control (or courage or energy) to carry out some declared intention of mine can I credibly give myself out as *someone who is going to act that way* ('Don't worry, I won't get into an argument about . . . '); if the condition is not met, others will do better to disregard my words in favour of whatever locally relevant knowledge they may have of my involvement in the 'realm of law' (say, the number of drinks, hours or minutes of dinner party, or whatever that it usually takes to crack my thin veneer of cool).[12]

The sovereign individual, as the positive pole of Nietzsche's contrast, refers to 'the condition of "self-mastery" or full competence to represent oneself to the rest of the world.'[13] At the negative pole of Nietzsche's contrast, it seems, stands 'the liar who breaks his word the moment he utters it', that is, in contemporary philosophical parlance: the wanton.[14] There is, I think, little doubt that Nietzsche draws this contrast in such extreme terms in order to heighten our attraction to the figure of the sovereign individual and our repulsion from the figure of the wanton, but in doing so he raises a puzzle to which Ridley has drawn attention, namely, what is distinctive about the sovereign individual's promise-making?[15] Since it is the case that the vast majority of socialized individuals are not wantons, that is, are capable of making and, *ceteris paribus*, keeping promises and since Nietzsche, as we have seen, spends some time in this essay explaining how this comes to be the case, what is it that distinguishes the sovereign individual?

In the first essay of the *Genealogy* (and elsewhere), Nietzsche ascribes to noble morality, and himself endorses, an account of agency in which one's

[11] Lovibond (2002: 71).

[12] Lovibond (2002: 72). These remarks preface a very interesting discussion of Nietzsche on the sovereign individual in which Lovibond develops a compelling account of being serious in uttering certain words as part of her overall reflections on ethical formation.

[13] Lovibond (2002: 74) It is notable that the ethical terms of this scale are provided by respect (in the appraisive rather than recognitive sense of this term, that is, as we might say, esteem) and contempt; terms that refer to the *character* of the agent. This is not only consistent with the form of noble morality in the first essay but also indicates that for those who understand themselves in the light of this ethical standpoint, the failure to sustain a commitment is a source of self-contempt.

[14] Frankfurt (1988: 11–25).

[15] Ridley (this volume).

deeds are seen as criterial of one's intentions, beliefs, desires, etc.[16] On this view, as Ridley points out, 'if it is essential to a promise's being made in good faith that the agent intend to act on it, it is essential, too, that—*ceteris paribus*—he does indeed so act.'[17] If, however, the figure of sovereign individual represents a self-conscious condition of self-mastery, this entails a specific kind of understanding of the *ceteris paribus* clause, that is, one in which the range of elaboratives to which one can have recourse is limited to reasons that are compatible with the presumption of self-mastery. There are thus two main types of excuse that could justify the failure to maintain a commitment which relate to conditions of causal and normative necessity respectively. The first is that honouring one's commitment is causally impossible due to circumstances beyond one's control; hence, one cannot *physically* do what is required (say, fly from London to New York today to be best man at a wedding since all flights are cancelled due to a terrorist attack). The second is that keeping one's promises is normatively impossible due to circumstances beyond one's control; hence, one must not *ethically* do what is required (say, ignore the drowning child in order to fulfil the obligation to meet a friend for a quiet drink and chat). Notice that a further implication of this self-understanding is that, even in circumstances where the reasons for breach of one's commitment are exculpatory, the sovereign individual acknowledges their accountability to the addressee of their commitment and, thus, an acknowledgment of the damage that may have resulted. This claim is supported by Nietzsche's characterization of the sovereign individual as 'anyone who promises like a sovereign [. . .] who is sparing with his trust, who *confers distinction* when he trusts, who gives his word as something which can be relied on, because he knows himself strong enough to uphold it even against accidents, even "against fate" '. The point here is not *per impossible* that the sovereign individual has (or is committed to) mastery over fate in general—a fantasy of which Nietzsche would be entirely dismissive—but that the sovereign individual is characterized by a degree of prudence in their commitment-making activity (that is, a serious effort to consider, as far as possible, the types of circumstance in which the commitment is to be honoured and the range of costs that may arise in fulfilment of the commitment as well as its prospects for conflicting with existing commitments), where this prudence is engendered precisely by an acknowledgment of one's responsibility as extending to those occasions on which the commitment cannot or must not be honoured in the form of an acknowledgement of the moral remainders that result from one's justified inability to keep one's word. Upholding one's word 'even "against fate" ' does not mean fantastically committing oneself to the incoherent goal of doing what is causally or ethically impossible for one to do, it means willingly acknowledging one's relationship to the damage incurred when one's commitment cannot or must not be kept. This is a stance which

[16] See Pippin (2004), Owen (2007), and Ridley (this volume). [17] Ibid., 4.

acknowledges and affirms the fatality of one's agency rather than seeking to avoid or deny it.

In relation to this first aspect of the distinctiveness of the sovereign individual, Nietzsche's position may be aligned with a point that Bernard Williams was wont to press against 'the morality system' whose standpoint he describes as granting no special significance to the thought *I did it* and hence, as turning 'our attention away from an important dimension of ethical experience, which lies in the distinction between what one has and what one has not done', a distinction that 'can be as important as the distinction between the voluntary and the non-voluntary'.[18] Williams' illustrates this point with the example of Ajax following his goddess-deranged slaughter of a flock of sheep:

> Ajax then wakes up and shows that he has recovered his mind. There is a passionate lyric outburst of despair and, above all, shame: he has made himself, apart from anything else, utterly absurd. It becomes increasingly clear to himself that he can only kill himself. He knows that he cannot change his *ethos*, his character, and he knows that after what he has done, this grotesque humiliation, he cannot lead the only kind of life his *ethos* demands. . . . Being what he is, he could not live as a man who had done these things; it would be merely impossible, in virtue of the relations between what he expects of the world and what the world expects of a man who expects that of it.[19]

Williams' point is not that we should endorse Ajax's suicide but that we should acknowledge the ethical intelligibility of Ajax's response and, hence, the weight that *I did it* can play in our ethical lives. For our immediate concerns, the point is this: the sovereign individual is one for whom the thought *I did it* has ethical purchase and salience.

There is, however, another dimension of the sovereign individual's promise-making that is distinctive. This second dimension also hangs on the expressive account of agency to which Nietzsche is committed and can be drawn out by contrasting promises whose success conditions (i.e., the conditions that entitle one to say that the promise has been kept) can and cannot be specified *externally* (i.e., in advance and independent of the execution of the accomplishment). If I promise to meet you today for lunch in the pub, the success conditions can be specified externally: I have kept my promise if I turn up at the pub in order to eat with you within the relevant time frame. By contrast, if I promise to love and honour you until death us do part, then what counts as keeping this promise cannot be fully specified in advance and independently of a particular way of keeping it. In the former case, keeping my promise simply confirms the presence of my intention; in the latter case, *the nature of my intention is revealed in the way that I keep it*. What is distinctive about the sovereign individual in this respect is that his most characteristic form of promise-making is of the latter type; indeed, it is precisely the sovereign individual's self-mastery that grants

[18] Williams (1985: 177) [19] Williams (1993: 72–3)

him the prerogative to engage in this kind of promise-making.[20] Another way of drawing the distinction between the two kinds of promise-making invoked here is to specify them in terms of commitments whose character is fully determined by the letter of the law and commitments whose character can only be fully determined by reference to both the letter and spirit of the law.[21] As Ridley comments, using the example of marriage:

> It is true that there are some independently specifiable success-conditions here (although they are defeasible). Respect is presumably necessary, for example, as are caring for the other person's interest and not betraying them, say. But what exactly might *count* as betrayal, or what caring for the other person's interests might *look* like in this case—or even whether *these* things are what is at issue—cannot be specified independently of the particular marriage that it is, of the circumstances, history, and personalities peculiar to it, and of how those things unfold or develop over time. It is, in other words, perfectly possible that everything I do is, as it were, strictly speaking respectful, considerate and loyal, and yet that I fail to be any good as a husband—I am true to the letter but miss the spirit, as we might say.[22]

This second aspect of the distinctiveness of the sovereign individual helps to illuminate three further points with respect to Nietzsche's argument.

First, it draws out the sense in which the sovereign individual can be represented by Nietzsche as the ethical *telos* of the process of socialization, which he is exploring in the second essay of the *Genealogy* since the freedom enjoyed and exemplified by the sovereign individual is only available to persons who are, in Ridley's phrase, 'socialized . . . all the way down', that is, individuals who have mastered the norms constitutive of the social practices and institutions in and through which they act (in this case, that of marriage).

Second, it offers an obvious line of response to a slightly curious argument by Acampora that occurs in the course of her rejection of the claim that the sovereign individual represents Nietzsche's account of ethical maturity. She writes:

> I can see how such an interpretation can be rendered consistent with Nietzsche's preoccupation with drawing distinctions between the herd and those who somehow escape it, but how could it be that the Nietzsche who so emphasizes *becoming*, and who is suspicious of the concept of the subject (as the 'doer behind the deed'), think it desirable—let alone *possible*—that a person could *ensure* his or her word in the future? How could one promise to do something, to stand security for something, that cannot be predicted and for which one is, in a sense, no longer the one who *could* be responsible for it?[23]

At first glance, this argument seems to verge on utter absurdity since some minimal capacity for making and, *ceteris paribus*, keeping commitments is a condition of being able to speak intelligibly of ethical agency at all. Perhaps

[20] Ridley (this volume). [21] Ridley (this volume). [22] Ridley (this volume).
[23] Acampora (2006: 153).

though Davis Acampora is struggling with a slightly different issue, namely, how to square Nietzsche's stress on the process of becoming with standing surety for one's word given that the latter, on her view, would have to imply that one takes oneself to be identical to the person who promised at the time that he or she promised? The mistake here is to fail to see that not only is promising compatible with becoming what one is, it is actually *essential* to it in two respects. First, it is not least in and through standing surety for one's commitments that one constitutes oneself as a self, as a being whose identity is characterized by continuity through time (note that this obviously need not imply some core identity that persists since it can equally be characterized by the image of a rope comprised of numerous overlapping threads of varying lengths or, indeed, of Theseus's ship in which all the parts are gradually replaced). Second, it is only through commitments characterized by internal success conditions that one discovers what one's intentions are or, put another way, acquires self-knowledge concerning what one is, what is necessary to oneself and what is not. In so far as the process of *becoming what one is* requires self-knowledge concerning what one is, then one's commitments and the issues that arise in the practical lived working out of what one has committed oneself to in making these commitments is of fundamental importance to this process.

The third point to note is that this second dimension of distinctiveness enables us to see once again that Nietzsche is articulating a view of ethical autonomy that contrasts sharply with the ideal of moral autonomy expressed in Kant and which Nietzsche takes to be representative of 'morality'. This is so because it directs attention to the fact that the central role of the categorical imperative in Kantian morality entails that if

> I find that the maxim of my action cannot be universalized without contradiction, I have identified an absolute prohibition, an unconditional 'I will not'. I have, in other words, stopped short at a formulable instruction that might be fully obeyed by anyone. . . . The spirit . . . has gone missing without trace.[24]

We can put the point like this: 'Morality' in the sense exemplified by Kant may have liberated itself from the morality of custom as regards to *content* but it has not done so with regard to *form*. Moral freedom for Kant, Nietzsche charges, can be articulated in terms of compliance with a list of 'I will not's that can be specified in advance and independently of the way in which commitment to them is executed. In this respect, Kant's philosophy exhibits the characteristic errors of 'morality', namely, a failure to acknowledge the expressive character of human agency combined with a stress on the unconditional character of moral imperatives, and does so in a way that leaves it blind to the nature and experience of human freedom as an unformulable process of self-legislation.

[24] Ridley (this volume).

It would, however, be a mistake to view Nietzsche's account of ethical autonomy purely and simply in terms of the picture of the *sovereign individual* since this picture addresses itself only to the issue of one's relationship to one's commitments as ends that are given and not as ends that are themselves open to reflective ethical scrutiny and assessment. As Robert Guay has cogently argued, for Nietzsche, freedom requires that we engage in critically distanced reflection on our current self-understanding. Nietzsche's point is that freedom demands 'the ability to take one's virtues and oneself as objects of reflection, assessment and possible transformation, so that one can determine who one is':

> As Nietzsche pointed out 'whoever reaches his ideal in doing so transcends it'. To take ourselves as potentially free requires that we are not merely bearers of good qualities but self-determining beings capable of distanced reflection. So to attain one's ideal is always that and also to attain a new standpoint, from which one can look beyond it to how to live one's life in the future.[25]

We can link these two aspects of freedom by noting how they fit naturally in Nietzsche's view of ethical education and self-transformation as a process of relating to and moving beyond exemplars conceived as concrete ideals, that is, individuals who have given a certain style to their characters and thus become able to serve as models, not for *imitation*, but for *following*.

What Nietzsche's position does share with Kant's moral philosophy is a view of autonomy as the ground of self-respect. Recall: 'The "free" man . . . possesses also in the property [the prerogative to promise] his *measure of value*: looking out at others from his own vantage point, he bestows respect or contempt. Necessarily, he respects those who are like himself—the strong and reliable (those with the *prerogative* to promise), that is, anyone who promises like a sovereign . . . ' As an exemplification of a key part of Nietzsche's account of ethical maturity, the sovereign individual marks out conditions of application of the concepts 'respect' and 'contempt' and their first person forms, namely, 'self-esteem (or pride)' and 'self-contempt (or shame)'.

We can add further support for the interpretation here advanced and give a final twist to it by attending to Nietzsche's discussion of freedom in *Twilight of the Idols*:

> *My conception of freedom.*—The value of a thing sometimes lies not in what one attains with it, but in what one pays for it—what it *costs* us. [. . .] For what is freedom? That one has the will to self-responsibility. That one preserves the distance that divides us. That one has become more indifferent to hardship, toil, privation, even to life. That one is ready to sacrifice men to one's cause, oneself not excepted. Freedom means that the manly instincts that delight in war and victory have gained mastery over the other instincts—for example, over the instinct for 'happiness'. . . . The free man is a

[25] Guay (2002: 315).

warrior.—How is freedom measured in individuals as in nations? By the resistance which has to be overcome, by the effort it costs to stay *aloft*. One would have to look for the highest type of free human beings wherever the highest resistance is constantly being overcome . . . . (TI *Expeditions* 38)

There are two arguments in this passage concerning the concept of freedom and the value of freedom respectively. The former endorses the view expressed in Nietzsche's remarks on the sovereign individual in identifying freedom with standing surety for oneself. The latter though draws attention to an issue that I have thus far left aside which is this: while the sovereign individual represents a condition of self-mastery, this condition cannot be seen, given Nietzsche's stress on our exposure to chance and necessity (i.e., luck), as a condition which can be fully and finally achieved. On the contrary, the condition of self-mastery must be construed as an ongoing achievement or, to put it another way, a continual process of struggle engaged with both the material of oneself and the circumstances of one's agency. Nietzsche is making two related psychological points on this latter issue. The first is that the degree to which one values freedom is likely to be dependent on the nature and extent of the obstacles that one has had to confront and overcome in standing surety for oneself. The second is that to the extent that one values freedom one will seek to test and extend one's powers. Taking these points together, it is clear that Nietzsche sees the sovereign individual, the ethically autonomous person, as one who is engaged in a process of struggle which continues through the course of one's life.

What, though, has any of this to do with self-love? Thus far it is has been established that Nietzsche is committed to an account of ethical autonomy understood in terms of binding one self to unformulable laws and that ethical autonomy is the ground of self-respect. In these respects, Nietzsche would seem, despite his criticisms of Kant, to be situated firmly with the party of *law* in the contest between love and law in European philosophy. To move beyond this impression of Nietzsche's position, we need to turn to the question of how his reflections on autonomy and self-respect are related to his account of self-love.

III

A starting point for discussion of Nietzsche and self-love is provided by section 334 of *The Gay Science*:

One must learn to love.—This happens to us in music: first one must *learn to hear* a figure and a melody at all, to detect and distinguish it, to isolate and delimit it as a life in itself; then one needs effort and good will to *stand* it despite its strangeness; patience with its appearance and expression, kindheartedness about its oddity. Finally comes a moment when we are *used* to it; when we expect it; when we sense that we'd miss it if it were missing; and now it continues relentlessly to compel and enchant us until we have become its humble and enraptured lovers, who no longer want anything better from the

world than it and it again. But this happens to us not only in music: it is in this way that we have *learned to love* everything that we now love. We are always rewarded in the end for our good will, our patience, our fair-mindedness and gentleness with what is strange, as it gradually casts off its veil and presents itself as a new and indescribable beauty. That is *its thanks* for our hospitality. Even he who loves himself will have learned it this way—there is no other way. Love, too, must be learned.

As this passage makes clear, Nietzsche, like Aristotle, takes self-love to be an acquired disposition. It is also the case that Nietzsche takes self-love to the natural disposition of the virtuous man; for example, he writes: '*The noble soul reveres itself*' (BGE 287). But if, as this remark suggests, Nietzsche follows Aristotle in formally identifying *true* self-love with the disposition to value what is noble, this immediately raises the question of the content of Nietzsche's conception of nobility. Two elements of Nietzsche's philosophy present themselves for reflection at this point. The first is the idea of *amor fati* offered in GS 276:

I want to learn more and more how to see what is necessary in things as what is beautiful in them. *Amor fati*: let that be my love from now on! I do not want to wage war against ugliness. I do not want to accuse; I do not even want to accuse the accusers. Let *looking away* be my only negation! And, all in all and on the whole: some day I want only to be a Yes-sayer!

Learning to love oneself is thus to be construed as learning to love the piece of fate that one is by coming to see what is necessary in and to oneself as 'a new and indescribable beauty'. The second is the doctrine of eternal recurrence set out in GS 341 according to which learning to love oneself is to be construed as coming to experience 'a tremendous moment' in which one is sufficiently well-disposed to oneself and to life to respond joyously to the demon's raising of the prospect of the eternal recurrence of one's life *exactly* as it has happened, that is, as a lover of one's life 'who no longer want anything better from the world than it and it again' (GS 334).[26] Nietzsche himself appears to take these elements as compatible in the sense of holding that being characterized by the disposition of *amor fati* is a sufficient condition for responding joyously to the prospect of eternal recurrence and *vice versa*. I am unsure that the two doctrines are compatible in the way that Nietzsche seems to envisage but, for the purposes of this essay, will set aside this issue by focusing on valuing the disposition of *amor fati* as the salient expression of Nietzsche's conception of *true* self-love.

We can begin exploring this thought by noting that GS 334 presents learning to love as a work of interpretation in which we attend to the object of love such as to come gradually to see it—or what is necessary in it—as a 'new and indescribable beauty'. In the case of self-love, this entails that learning to love one's self is to engage in an artistic work of interpretation on oneself. Thus, an initial point to note is that Nietzsche's presentation of the acquisition of self-love

[26] Or perhaps to recognize oneself as having experienced such a moment . . . .

as a form of artistic interpretation through which one comes to see oneself as a 'new and indescribable beauty' in GS 334 refers us both forward to the following section of *The Gay Science* in which Nietzsche exhorts us to the ethical work needed to '*become what we are*—human beings who are new, unique, incomparable, who give themselves laws, who create themselves!' (GS 335) and backward to GS 290 in which the achievement of self-love is linked to that of giving style to one's character. Working through the relationship of Nietzsche's reflections on learning to love to these passages will provide us with the basis for understanding self-love as valuing the disposition of *amor fati*.

The pertinent remarks in the section 290 '*One thing is needful*' read as follows:

> To 'give style' to one's character—a great and rare art! It is practised by those who survey all the strengths and weaknesses that their nature has to offer and then fit them into an artistic plan until each appears as art and reason and even weaknesses delight the eye. Here a great mass of second nature has been added; there a piece of first nature removed—both times through long practice and daily work at it. Here the ugly that could not be removed is concealed; there it is reinterpreted into sublimity. Much that is vague and resisted shaping has been saved and employed for distant views—it is supposed to beckon towards the remote and immense. In the end, when the work is complete, it becomes clear how it was the force of a single taste that ruled and shaped everything great and small—whether the taste was good or bad means less than one may think; it's enough that it was one taste! It will be the strong and domineering nature who experience their most exquisite pleasure under such coercion, in being bound by but also perfected under their own law; . . . Conversely, it is the weak characters with no power over themselves who *hate* the constraint of style: they feel if this bitterly evil compulsion were to be imposed on them, they would have to become *commonplace* under it—they become slaves as soon as they serve; they hate to serve. (GS 290)

Before turning to the main issues arising from this passage, it is worth noting—and hence warding off—a potential misunderstanding. The misunderstanding would consist of taking the opening of this passage to encourage the thought that giving style to one's character is accomplished by identifying the materials, devising a blueprint and then simply executing the plan. The misunderstanding here is to mistake the appeal to an 'artistic plan' as if the movement from plan to work was simply the prescriptive activity of applying a rule given in advance rather than—as, say, the image of movement from preliminary sketch to actual painting suggests—the expressive activity of giving a rule in and through the performance. It is for this reason that it is only when the work is complete that the operation of a single taste can be discerned since the taste is constituted as the singular taste that it is in and through the performance.

It is clear then that Nietzsche is describing a process of self-creation (or, perhaps better, self-formation) characterized by the activity of submitting to what Nietzsche refers to, under the aesthetic aspect, as 'one's taste' and, under the ethical aspect, as one's 'own law'. But what constitutes one's taste or one's own

law? What is one submitting to here? Is one submitting in any intelligible sense at all? This last question arises because on one understanding of the ideas of one's taste or one's own law, these are simply whatever one chooses, one's inclinations, at any given moment and hence the idea of submission loses intelligible application. On such a view, the distinction—crucial to Nietzsche—between strong and weak collapses. To distinguish Nietzsche's position from this view—and thereby clarify what it is to possess, exercise, and submit to one's taste—it is helpful to recall the interpretive point made above (and stressed throughout this essay) that Nietzsche is committed to an expressive view of agency exemplified by the relation of the artist to his work, a relation that he characterizes thus:

> Every artist knows how far from any feeling of letting himself go his most 'natural' state is—the free ordering, placing, disposing, giving form in the moment of 'inspiration'—and how strictly and subtly he obeys thousandfold laws precisely then, laws that precisely on account of their hardness and determination defy all formulation through concepts . . . . (BGE 188).

The salient point for Nietzsche, as Aaron Ridley notes, is that artistic activity

> is law-like, in the sense that it is possible to go wrong, to make mistakes. Yet the laws against which these mistakes offend often declare themselves only in the moment at which they are breached, indeed *in* the breaching of them. And this is why getting something *right* feels like—is—getting what one was after all along, even when one could not have said in advance precisely what that was. In this way, successful artistry is also a form of self-discovery—it is the discovery, in the lawfulness of one's actions, of the innermost character of one's intentions: . . .[27]

To submit to one's taste or one's own law is, thus, to be bound by constraints or, more precisely, to bind oneself to constraints that, at least in advance, 'defy all formulation through concepts'—and this is *to affirm* such constraints as the conditions of one's agency. But what generates these constraints that the artist affirms as the conditions of his agency? The answer here—and hence the response to the question of what comprises one's taste or one's own law—is the intentions of the artist which, in the case of self-artistry, are comprised of one's evaluative commitments.[28] Thus far, then, Nietzsche may be read as

[27] Ridley (2005: xv); see also Ridley (2006).

[28] Recognizing this point, it is worth stressing here that the fact that Nietzsche speaks of the constraint of style in terms of submitting to the demands of one's own taste or law does not mean that the relevant constraints are not objective in the sense of being independent of the person concerned. Consider that a commitment to the value of friendship may be an expression of one's taste, but, if this is the case, then the constraints to which one is subject on account of one's taste are those norms and necessities that govern the practice of friendship. Similarly, a commitment to the value of scientific activity entails that the constraints to which one is subject are those norms and necessities that govern the practice of scientific activity. Hence it follows that one can only give style to one's character by submitting to the norms and necessities that govern the relationships, practices, ideals, etc. to which one is committed.

offering a re-description of the sovereign individual from an aesthetic standpoint. A preliminary view concerning the relationship between self-love and self-respect would thus be that valuing the disposition of *amor fati* broadly equates to valuing the will to self-responsibility (i.e., ethical autonomy) and, hence, that self-love consists in valuing self-respect. This is, I think, broadly the right view but we have not yet uncovered the whole story.

We can move forward by recalling that one's intentions are only discovered through the work on oneself and taking up the necessary correlate of this point, namely, that the expression of these intentions (and hence what they are) is not independent of the medium of expression, the material through which the artist seeks to realize his intentions, and hence of the constraints imposed by this material. In the case of self-artistry, this medium is the human material—nature and second nature—comprising oneself ('Here a great mass of second nature has been added; there a piece of first nature removed—[. . .]. Here the ugly that could not be removed is concealed; there it is reinterpreted into sublimity. Much that is vague and resisted shaping has been saved and employed for distant views [. . .].') Recognizing the salience of the issue of the medium of self-artistry helps not only to explain why Nietzsche's remarks on giving style to one's character begin by stressing that those who engage in this art 'survey all the strengths and weaknesses that their nature has to offer', but also provides the link between the concluding reference to self-love in GS 334 and the discussion of intellectual conscience that immediately follows in the section '*Long live physics*!'. In GS 335, Nietzsche sets out an account of intellectual conscience in terms of a commitment to truthfulness (*Redlichkeit*) oriented to becoming 'the best students and discoverers of everything lawful and necessary in the world' and, specifically, what is necessary in us as nature and, more particularly, as second nature. The role of intellectual conscience is precisely to seek to determine what is necessary and what is not in us, that is, to work out what are the constraints that necessarily attend expression through the medium of human material in general and of oneself in particular; its role is to free us from both the illusion of false necessities (as produced by, for example, morality) and an idealism that fails to acknowledge necessary constraints on our agency.

The exercise of intellectual conscience is, thus, integral to the process of self-artistry and this fact has particular significance for establishing the relationship between the process of self-artistry described in GS 290 and the account of *amor fati* offered in GS 276. Recall that in GS 290 Nietzsche distinguishes between non-necessary weaknesses that can be removed or overcome and necessary weakness that are to be concealed or reinterpreted and made sublime. These two features of the process of self-artistry are directly connected to the statement of *amor fati* in GS 276. As Ridley comments:

The connection comes to this: becoming what you are requires that you distinguish between what is and what is not necessary in things, including yourself (a job for the intellectual conscience). What is not necessary, and is weak or ugly, should be removed. What *is* necessary should, if weak or ugly, either be concealed ('*Looking away* shall be my only negation' (GS 276)) or else 'reinterpreted', so that one learns how to see it as beautiful, as a strength.[29]

The implication of this point for my argument is clear: *true* self-love as valuing the disposition of *amor fati* entails engaging in the process of self-artistry that Nietzsche describes in GS 290—a conclusion that explains the title of that aphorism: '*One thing is needful*'. But if this is the case, it provides us with a way of understanding the relationship between self-respect and self-love in Nietzsche's ethics that is closely related to the issue of the relationship of truthfulness and artistry in his philosophy as a whole.

Consider that a necessary condition of the self-mastery exemplified by the sovereign individual is the exercise of intellectual conscience and, hence, truthfulness (*Redlichkeit*) with respect to what one is, what one is committed to and the requirements of these commitments. At its limit, this is an unflinchingly honest and realistic appraisal of oneself and the circumstances of one's agency; it is the type of appraisal that Nietzsche so admired in Thucydides' account of the Peloponnesian War and in Machiavelli's reflections on the circumstances of political agency in Renaissance Italy. However, Nietzsche is well aware of the difficulty of exhibiting or sustaining such an unflinching stance (hence his acknowledgment that how much truth you can bear is a measure of strength). More generally, Nietzsche acknowledges that the potentially ruthless truthfulness of intellectual conscience requires a counterforce if it is not to lead to nausea and suicide, where this counterforce is 'art, as the *good* will to appearance' (GS 107). Nietzsche comments: 'As an aesthetic phenomenon, existence is still *bearable* to us, and art furnishes us with the eye and the hand and above all the good conscience to be *able* to make such a phenomenon of ourselves' (GS 107). Hence, while the ethical autonomy that grounds self-respect requires that one exhibit the virtue of truthfulness, a condition of being able to maintain the kind of courageous commitment to engagement in the world exemplified by the sovereign individual is that one has recourse to art when it is necessary.

Consider the case of Nietzsche himself as an illustration of this point. In *The Antichrist*, Nietzsche remarks of the service of truth to which he has committed himself:

Truth has to be fought for every step of the way, almost everything else dear to our hearts, on which our love and trust in life depend, has had to be sacrificed to it. Greatness of soul is needed for it: the service of truth is the hardest service.—For what does it mean

[29] Ridely (2005: xiii).

to be *honest* in intellectual things? That one is stern towards one's heart, that one despises 'fine feelings', that one makes every Yes and No a question of conscience! (A 50)

A clear illustration of the recourse to art that may be necessary to maintaining this overall commitment to truth is provided by Nietzsche's remarkable reflections on himself in the 1886 Preface to *Human, All Too Human*:

My writings have been called a schooling in suspicion, even more in contempt, but fortunately also in courage, indeed in audacity. And in fact I myself do not believe that anyone else has ever looked into the world with an equally profound degree of suspicion [. . .] and anyone who could divine something of the consequences that lie in that profound suspiciousness, something of the fears and frost of isolation to which that unconditional *disparity of view* condemns him who is infected with it, will also understand how often, in an effort to recover from myself, as it were to induce a temporary self-forgetting, I have sought shelter in this or that—in some piece of admiration or enmity or scientificality or frivolity or stupidity; and why, where I could not find what I *needed*, I had artificially to enforce, falsify and invent a suitable fiction for myself (—and what else have poets ever done? and to what end does art exist in the world at all?). What I again and again needed most for my cure and self-restoration, however, was the belief that I was *not* thus isolated, not alone in *seeing* as I did—an enchanted surmising of relatedness and identity in eye and desires, a reposing in a trust of friendship, a blindness in concert with another without suspicion or question-marks, [. . .] Perhaps in this regard I might be reproached with having employed a certain amount of 'art', a certain amount of false-coinage: for example, that I knowingly-willingly closed my eyes before Schopenhauer's blind will to morality at a time when I was already sufficiently clear-sighted about morality; likewise that I deceived myself over Richard Wagner's incurable romanticism, as though it were a beginning and not an end; likewise over the Greeks, likewise over the Germans and their future—and perhaps a whole long list could be made of such likewises?—Supposing, however, that all this were true and that I was reproached with it with good reason, what do *you* know, what *could* you know, of how much cunning in self-preservation, how much reason and higher safeguarding, is contained in such self-deception—or of how much falsity I shall *require* if I am to continue to permit myself the luxury of *my* truthfulness? (HAH P 1)

This is an astonishing passage not least as an exercise of intellectual conscience that acknowledges that Nietzsche's own global commitment to truth has necessitated local falsifications, projections, rationalizations, and illusions. For our current purposes though its central significance is that it serves as an exemplification of Nietzsche's understanding of self-love conceived as valuing the disposition of *amor fati*; it shows that valuing this disposition is not only a matter of valuing truthfulness as the good will to reality, but also, and crucially, of valuing artistry as the *good* will to appearance in so far as it is *necessary* to supporting one's active immersion in life, to maintaining the will to self-responsibility in and through which one affirms the fatality of one's agency.[30]

[30] Taking the appearance/reality distinction is its thorough un-superlative non-metaphysical sense.

The upshot of our discussion is thus that Nietzsche seeks to dissolve the tension between self-respect and self-love by presenting the latter as the disposition to value the former. This returns us to the question of the nature of Nietzsche's objection to the distinction between morality and egoism in modern moral philosophy, but before addressing this point, let us briefly situate Nietzsche's position in relation to Aristotle and Kant. First, he joins both Aristotle and Kant in rejecting the life of pleasurable inclinations, of pseudo self-love. I haven't stressed this point but it is obviously apparent from the discussion of freedom in TI *Expeditions* 38 (see also 39 and 41) and the contrast between strong and weak drawn in GS 290. Second, he adopts Aristotle's formal notion of *true* self-love as valuing what is noble. Third, he offers an account of nobility that aligns him with Kant (and Hegel) in stressing the centrality of autonomy as the principle of modern ethics. Fourth, his account of autonomy makes appeal not to universal moral rules but to the exercise of virtues. What light does this shed on Nietzsche's objection to the modern morality/egoism distinction at least as it is exemplified in Kant's moral philosophy?

At an initial level, Nietzsche's objection is simply that Kant's framing of this distinction rules out self-love as the general orientation to *eudaemonia* (whether conceived in Aristotelian or Nietzschean terms) and, hence, fails to acknowledge the centrality of self-love to the achievement and maintenance of the disposition of *amor fati* (Nietzsche's version of eudaemonia). Of course, this failure on Kant's part has, in Nietzsche's view, deep roots in the ascetic ideal, but, for our purposes, a fairly economical way of making clear what motivates Nietzsche's assertion of the significance of self-love is to recall that Kant's rejection of self-love is bound up with his effort to immunize morality from luck and to note that Nietzsche cannot possibly endorse this project. First, Nietzsche's commitment to naturalism in ethics is understood by him to entail acceptance of the claim that our exposure to chance and necessity goes all the way down—and hence that the project of seeking to immunize morality from luck is a non-starter. Second, Nietzsche holds that the illusion of intelligibility that characterizes this project is dependent on a metaphysical version of the appearance/reality distinction that is incoherent and betrays a hostility to the conditions of human life. Third, the project requires an erroneous understanding of the idea of the moral law as necessarily universal that is predicated on abstracting from the agent and the circumstances of their agency to an imaginary idealized and de-contextualized agent in a way that undermines and obscures the character of ethical autonomy. Hence, Nietzsche's strategy is exactly the opposite of Kant's; it is to build an acknowledgment and affirmation of our exposure to luck into our understanding of ethics and ethical autonomy. This strategy embraces self-love as necessary to the task of developing and maintaining an orientation to, and engagement with, the world that consists in *leading* our lives and not simply existing through the course of their duration.

CONCLUSION

If the argument offered in this essay is cogent, perhaps its most striking finding is that it would appear that Nietzsche has a rather greater claim than Kant to have resolved the tension between love and law that is characteristic of European moral philosophy. At the same time, though, this finding is dependent on my arguments for two further highly controversial theses, namely, that Nietzsche is committed to ethical autonomy as the basic principle of modern ethics and that there is a fundamental connection between ethical autonomy and the disposition of *amor fati*. Not the least of the consequences that follow if these arguments are compelling is that Nietzsche may have rather more to contribute to debates in contemporary ethics between neoKantians and neoAristotelians than either party may currently wish to admit.[31]

REFERENCES

ABBREVIATIONS

Aristotle

NE Nichomachean Ethics, trans. D. Ross, revised J. L. Ackrill & J.O. Urmson (Oxford: Oxford University Press, 1980)

Kant

AP Anthropology from a Pragmatic Point of View, trans. V.L. Dowdell (Carbondale: Southern Illinois University Press, 1978).

CPR Critique of Practical Reason, trans. M. J. Gregor, in *Immanuel Kant: Practical Philosophy* (Cambridge: Cambridge University Press), 133–271.

G Groundwork of the Metaphysics of Morals, trans. M. J. Gregor, in *Immanuel Kant: Practical Philosophy* (Cambridge: Cambridge University Press), 37–108.

MM The Metaphysics of Morals, trans. M. J. Gregor, in *Immanuel Kant: Practical Philosophy* (Cambridge: Cambridge University Press), 353–603.

WRITINGS BY NIETZSCHE

The Antichrist, trans. R. J. Hollingdale (Harmondsworth, Penguin, 1968).

Beyond Good and Evil, ed. R-P. Horstmann and J. Norman (Cambridge, Cambridge University Press, 2002).

Daybreak, ed. M. Clark and B. Leiter, trans. R. J. Hollingdale (Cambridge, Cambridge University Press, 1997).

[31] I do, of course, except Christine Swanton from the scope of the remark since she has, from a different direction, been exploring this issue sensitively for some years.

On the Genealogy of Morals, trans. D. Smith (Oxford, Oxford University Press, 1996).
The Gay Science, ed. B. Williams, trans. Josephine Nauckhoff and Adrian del Caro Cambridge, (Cambridge University Press, 2001).
Human, All Too Human, ed. E. Heller, trans. R. J. Hollingdale (Cambridge, Cambridge University Press, 1986).
Twilight of the Idols, trans. D. Large (Oxford, Oxford University Press, 1998).

OTHER WORKS CITED

Acampora, Christa Davis (2006). 'On Sovereignty and Overhumanity: Why it matters how we read Nietzsche *Genealogy* II: 2', in Christa Davis Acampora (ed.), *Nietzsche's Genealogy of Morals: Critical Essays* (Lahnam: Rowman & Littlefield), 147–62.
Frankfurt, Harry (1988). *The Importance of What We Care About* (Cambridge: Cambridge University Press).
Gemes, Ken (2009). 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual' (this volume).
Guay, Robert (2002). 'Nietzsche on Freedom', *European Journal of Philosophy* 10: 302–27.
Hatab, Lawrence (1995). *A Nietzschean Defense of Democracy* (Chicago, Ill: Open Court).
Lovibond, Sabina (2002). *Ethical Formation* (Cambridge, Mass.: Harvard University Press).
Owen, David (2007). *Nietzsche's Genealogy of Morality* (Stockfield: Acumen Press).
Pippin, Robert (2004). 'Lightning and Flash, Agent and Deed (GM I: 6–17)' in O. Höffe (ed.), *Friedrich Nietzsche, Genealogie der Moral* (Berlin: Academie Verlag), 47–63.
Ridley, Aaron (2005). 'Introduction', in Aaron Ridley and Judith Norman (eds.), *The Anti-Christ, Ecce Homo, Twilight of the Idols and other writings* (Cambridge: Cambridge University Press).
____ (2006). 'Nietzsche on Art and Freedom', *European Journal of Philosophy* 15: 2, 204–24.
____ (2009). 'Nietzsche's Intention: What the Sovereign Individual Promises' (this volume).
Schneewind, J.B. (1998). *The Invention of Autonomy* (Cambridge: Cambridge University Press).
Williams, Bernard (1985). *Ethics and the Limits of Philosophy* (London: Fontana Books).
____ (1993). *Shame and Necessity* (Berkeley: University of California Press).

This page intentionally left blank

11

The Eternal Recurrence: A Freudian Look at What Nietzsche Took to Be His Greatest Insight

Mathias Risse

1. INTRODUCTION

1.1 Nietzsche is a naturalist at least in the sense that much of his mature thought approaches philosophical problems in a manner inspired by the emerging biological sciences.[1] Characterizing his *Genealogy* as an exercise in "animal psychology" (GM III 20), he tries to explain in this way the development of a form of inner life shaped by *ressentiment* and guilt as well as by the appeal of ideals of self-denial. The *Genealogy* is guided by the idea that "our thoughts, values, every 'yes', 'no', 'if' and 'but' grow from us with the same necessity with which a tree bears its fruits—all related, and each with an affinity to each, and testimony to one will, one health, one earth, one sun" (*Preface*, 2). That is, the endorsement of values is to be explained like other behavior. As we read in *Beyond Good and Evil*, "[m]oralities are the sign languages of affects" (BGE 187), or in *Dawn*, "our moral judgments and evaluations (. . .) are only images and fantasies based on a physiological process unknown to us" (D 119). Morality and religion disregard this supervenience-relationship between psycho-physical type-facts and facts about beliefs and values. That is why Nietzsche calls them the realm of imaginary causes and effects (A 15).

[1] For references to and abbreviations of Nietzsche's and Freud's works, see the literature references. I am grateful to audiences at the Pacific APA in Oregon in March 2006 (in particular to John Richardson, who commented on this paper on that occasion), at the conference on "Nietzsche on Self, Agency, and Autonomy" at Birkbeck College, London, in April 2006, at the workshop on "Nietzsche and Morality" at Princeton University in November 2006, and at a workshop at the University of Texas in Austin in February 2007 (especially to Christopher Raymond, who commented on that occasion) for useful discussions. Thanks to Robert Pippin for suggesting the formulation "affirmation without affirmability."

Since there is a richness of such types, a universal "Thou Shalt" is misguided (TI V 6).[2]

Setting aside other concerns about the naturalist reading, this study explores its compatibility with the doctrine of the *Eternal Recurrence* (ER). That doctrine is most plausibly seen as a life-affirmation test: asking if one would live one's life again and again means testing if one has doubts about what kind of person one is. Nietzsche calls ER "the highest formulation of affirmation attainable" (EH III Z-1; see BGE 56), and himself its teacher (TI X 5). He takes ER to be his greatest insight, and "[t]he body of philosophically sophisticated and interesting criticism of [ER] remains unmatched by the literature on any other aspect" of his philosophy (Clark 1990: 245).

A number of naturalist readings do not make much of ER. Leiter (2002: 119, 287) briefly discusses it in his account of the higher men, and, again briefly, returns to it when discussing ascetic ideals. Richardson (2004: 197) mentions it in the context of Nietzsche's views on breeding exceptional individuals. Exploring Nietzsche's naturalism as an historian, Moore (2002) does not mention ER. So commentators championing the naturalist reading stand apart from those who regard ER as Nietzsche's central thought (Hatab 2005; Löwith 1978; Magnus 1978; Stambaugh 1972), or grant it a central role (Nehamas 1985; Clark 1990; Danto 1965). One must wonder whether this approach emphasizes components of his thought that are hard to reconcile with his commitment to ER, components that may receive less attention if one reads him along existentialist, transcendental, or "post-modern" lines. This study seeks to explore what a naturalist reading can make of Nietzsche's insistence on the importance of ER.[3]

1.2 Enters Freud. A whole literature is devoted to similarities between Nietzsche and Freud, primarily from a Freudian standpoint.[4] These similarities register strongly on a naturalist reading of Nietzsche because that reading emphasizes areas of overlap between them. Yet Freud never engaged with ER, nor did he even have a sense of its importance for Nietzsche.[5] One needs to wonder whether

[2] For the language of type-facts, see Leiter (2002), pp. 8–10. Strictly speaking, Nietzsche only refers to the *Genealogy's* second treatise as an exercise in animal psychology; but plausibly, if that treatise is of this genre, so are the other two and thus the work as a whole.

[3] Those components, of course, are part of his thought, regardless of how one characterizes his work overall. So one must ask at any rate about the coherence of his insistence on ER with these components.

[4] These investigations have also looked at references to Nietzsche that Freud made throughout his life to assess whether he may have gotten ideas from Nietzsche without giving proper credit; see the scattered remarks in Jones (1953), as well as Assoun (2000); Chapman and Chapman-Santana (1995); Greer (2002); Scavio et al. (1993); Gödde (1999); Lehrer (1995); and Gasser (1997). Their findings diverge widely, but Gasser, the most thorough of them, concludes that that is not the case.

[5] Freud (1920: 22), quotes the phrase "perpetual recurrence of the same thing" in his argument for the death drive. Some of the literature on the relationship between Freud and Nietzsche picks up on this "connection" between that argument (which turns on the postulation of a compulsion to repeat) and ER, as well as on the importance of "repeating" for Freud's understanding of neuroses;

Freud's own naturalist commitments discouraged him from exploring ER. If so, one needs to ask whether whatever discouraged Freud should have discouraged Nietzsche too. While this is a good reason for exploring why Freud had no interest in ER, one might wonder whether exploring Freudian themes is not a detour as far as the compatibility of an endorsement of ER with a naturalist reading of *Nietzsche* is concerned. It is, however, in the nature of their similarities that, frequently, Freud provides laborious investigations where Nietzsche offers flashy insights that require such elaboration to be taken seriously.[6] By drawing on Freud we can assess the concern that an endorsement of ER might well be consistent with other views Nietzsche endorses only because he does not say much about them. So exploring why Freud would have rejected ER is a fruitful way of assessing what to make of ER within a naturalist reading of Nietzsche.[7]

I first discuss ER through the *Gay Science*. Next I discuss three grounds on which Freud would reject ER. First, Freud takes an epistemological stance that leaves no room for the question ER answers. On this point Nietzsche disagrees, and for Freud of course it is irrelevant that a doctrine for which he has no use sits uneasily with his other views, but there are other reasons why Freud would have rejected ER that draw on views Nietzsche shares. The second reason turns on Freud's account of the mind, which accommodates some of Nietzsche's views. According to this model it is hard to know oneself well enough to be confident that one wants to live this life again and again. The third reason, finally, turns on Freud's view of guilt. Unlike Nietzsche, Freud does not think guilt can be overcome. There will always be this element in a person's life that she would want to be without. Upon inspection, Nietzsche should be closer to Freud on this point than it appears. Therefore this Freudian reason for resisting ER as a life-affirmation test also applies to him. In concluding I consider several possible resolutions of this tension in which his greatest insight stands with other elements of his thought. All are based on the observation that we find none of Freud's or Nietzsche's views to be actually *inconsistent* with ER. Yet none of them will be entirely satisfactory.

see Freud (1914), and Jones (1953: vol. 3, 271) (who, to his credit, does not make much of the fact that Freud quotes that phrase in Freud 1920); Assoun (2000: 178–81); and Chapelle (1999). Yet with a proper understanding of ER at hand, we can see that pursuing such parallels is not fruitful. Gasser (1997: 389–97) is aware of that.

[6] Examples include the theory of dreams, the notion of the unconscious, the origin and meaning of the term "id" in Freud's mature theory of the mind, the introduction of the phenomenon of repression, and the development of guilt. For details, see in particular Gasser (1997); Lehrer (1995); and Assoun (2000).

[7] There may be other reasons why Freud would reject ER. For instance, he may have thought it would be narcissistic to express this sort of satisfaction with one's life, or that a concern with an infinite repetition of one's life might testify to a rather troubled relationship with one's own past. But since my concern is with Nietzsche (and more specifically, with Nietzsche's philosophy, rather than his state of mind as psychoanalysis might see it), I am interested in Freudian reasons only to the extent that they are useful for understanding Nietzsche.

2. THE ETERNAL RECURRENCE

2.1 Any assessment of what to make of ER in the context of a naturalist reading depends on how we interpret ER. Following Hatab (2005), we can distinguish the following accounts: the cosmological reading, according to which ER is a quasi-scientific theory of cosmic time; the existential reading, which takes ER to be a test for life-affirmation; the decision-test reading, according to which ER functions parallel to the Categorical Imperative, evaluating actions in terms of the desirability of their eternal repetition; and symbolic readings, which take ER to be concerned with something other than the repetition of life.[8] I will assume, rather than argue, that we should adopt the existential reading. Other subscribers to (version of) this view include Magnus (1978); Clark (1990); and Nehamas (1985). The published writings in which ER appears are the *Gay Science, Zarathustra, Twilight of Idols,* and *Ecce Homo. Gay Science* IV provides the best understanding of ER, and it supports the existential reading. So I will introduce ER through GS IV.[9]

The *Gay Science* reflects on the "death of God" announced in GS 125, in Book III. Nietzsche's point is not to proclaim that "God is dead", that is, that the belief in the God of Christianity has become incredible (GS 343, GS 357, GS 377). He takes this for granted: the madman in GS 125 speaks to atheists. Yet the implications of this development are poorly understood. As GS 108 says, inaugurating GS I the shadow of Buddha was still shown in a cave centuries after his death. This image strikingly assesses the magnitude of Nietzsche's concerns: not only was that shadow still around centuries later, it was shown at a location without natural light. With the death of God, all corollary constructs must fall too (TI IX 5). Realizing this is more momentous than parting from the notion of God itself. Book IV offers an alternative to the belief in God, captured by ER and presented in GS 341. GS is committed to de-deification and naturalization (GS 109; cf. GS 346).

Just as Book III is structured around GS 125, so Book IV is built around GS 341, its penultimate section.[10] Book IV begins by introducing the idea of

[8] Hatab (2005) adds a fifth category, ontological readings. But these are specific ways of filling in what is meant by the "symbolic" content of ER as explicated by the fourth reading.

[9] This view conflicts with Clark (1990), who says *Zarathustra* is "the crucial text for determining which doctrine(s) of recurrence Nietzsche actually taught. He tells us that eternal recurrence is the 'fundamental conception' of *Zarathustra*, and, except for GS 34, provides no other independent basis for interpreting it" (p. 256). I do not discuss two major philosophical questions about this reading: whether it turns on there being a possibility of such infinite repetition to come true; and whether one needs to affirm the repetition of everything, including cruel episodes of history, to pass this test. The only point that matters for my argument is that the existential reading is the preferred one, and that view is at any rate widely shared in the literature.

[10] Book IV was originally the last book of GS. So GS 341 was originally the penultimate passage of the whole work, followed only by a reference to Zarathustra, and hence to Nietzsche's next

personal providence. As Nietzsche says in GS 276 (first section of Book IV): "I will ever more try to regard the necessary in things as the beautiful and thus become one of those who make things beautiful. Amor fati—let that be my love from now on." What at first looks like an expression of optimism on the occasion of the New Year (1882) is the core idea of his proposal in response to the "death of God". He explains it in more detail in GS 277.

Sittliche Weltordnungen (A 26), metaphysical-ethical worldviews, have become incredible, and Nietzsche's primary interest is in Christianity. Within these schemes of divine creation, the creator-god provided a fate, a providence (GS 277) for each person. Moreover, various senses in which one might ask about the meaning of life were intelligible within such frameworks: there were answers to questions such "why do I exist", or "what is the purpose of my life". What is required after such constructs have become incredible, and what Nietzsche intends to endorse cheerfully, is that each person devise her own providence: "to leave the gods alone (. . .) and to make do with the assumption that our own practical and theoretical dexterity in interpreting and understanding events has reached its zenith" (GS 277). This theme reappears in GS 290, where he points out that it is crucial to "give style" to one's character. Each person ought to aspire at being "satisfied with himself"; otherwise one will want to make others pay for it. GS 294 provides an alternative understanding of what it is to be satisfied with oneself: not to be afraid of oneself. The theme of "giving style" to oneself is resumed in GS 299, where Nietzsche inserts one of his most beautiful formulations, encouraging us to become "poets of your lives". GS 301 revisits that theme, and Nietzsche explains that "we have created the world that concerns human beings". In particular values, according to Nietzsche, are not provided by a metaphysical-ethical worldview, but created by us (cf. also GS 335).

2.2 Nietzsche's vision of personal providence culminates in GS 341, the first published mention of ER (GS 285 alludes to it). GS 340 introduces Socrates' last words: "Oh Crito, I owe Asklepsios a rooster". Asklepsios is the god of health, and the Greek custom was to sacrifice a rooster to him in gratitude for being cured. The fact that the dying Socrates owns up to this debt reveals that he saw life as a disease. This is striking because, as Nietzsche maintained since the *Birth of Tragedy*, Socrates is responsible for the emergence of metaphysical-ethical worldviews. Nietzsche declares the failure of that project through its originator before inaugurating his alternative. The next passage is GS 341:

The greatest weight.—What, if some day or night a demon were to steal after you in your loneliest loneliness and say to you: "This life as you now live it and have lived it, you will have to live once more and innumerable times more; and there will be nothing new in it, but every pain and every joy and every thought and sigh and everything

book, a reference that seems to be meant as a pointer that the theme of ER is further developed in that work.

unutterably small or great in your life will have to return to you, all in the same succession and sequence—even this spider and this moonlight between the trees, and even this moment and I myself. The eternal hourglass of existence is turned upside down again and again—and you with it, speck of dust!"—Would you not throw yourself down and gnash your teeth and curse the demon who spoke thus? Or have you once experienced a tremendous moment when you would have answered him: "You are a god and never have I heard anything more divine!" If this thought gained possession of you, it would change you as you are or perhaps crush you; the question in each and every thing, "Do you desire this once more, and innumerable times more?" would lie upon your actions as the greatest weight! Or how well disposed would you have to become to yourself and to life *to crave nothing more fervently* than this ultimate eternal confirmation and seal?—

What is left to individuals after the collapse of metaphysical-ethical worldviews is to create their own providence. As GS 124 says, we are now "at sea". The only way of obtaining a meaning is through affirming our own lives as they unfold in *this* world. Individuals affirm their lives fully only if they would want to repeat this life infinitely often, thus assuming "the greatest weight". GS 115 tells us that in the past we committed four major errors: we saw ourselves as incomplete and, as we may read this, felt compelled to see ourselves under the description of creatures of God; we added fictitious features to ourselves, which we may take to mean the Christian *sittliche Weltordnung*; we saw ourselves in a misguided rank order vis-à-vis animals; and we introduced new tables of values, considering them eternal and unconditional. As Nietzsche tells us in EH I in his reflection on GM, ascetic ideals, which motivated all these errors, succeeded because there was no alternative that could provide meaningful lives. ER is designed to do that, and thus to fill the gap opened by the decline of *sittliche Weltordnungen* in the only credible way.

3. THE SCIENTIFIC WORLDVIEW

3.1 Let us turn to the grounds on which Freud would have rejected ER. The first reason draws on epistemological commitments. Like Nietzsche, Freud seeks to explain the origins of moral beliefs and emotions. Unlike Nietzsche, he does not thereby aim to debunk morality. Neither does he reject common morality on behalf of human excellence, nor does he wish to replace debunked ideals with new ones. While the *Future of an Illusion* is clear that "[o]ur God, Logos, promises no compensation for us who suffer grievously from life" (Freud 1927: 54), Freud thought morality can be justified on self-interested grounds (cf. Freud 1927: 41). However, as he reiterates in *Civilization and Its Discontents,*

[t]he question of the purpose of human life has been raised countless times; it has never yet received a satisfactory answer and perhaps does not admit of one. (. . .) It looks (. . .) as though one had a right to dismiss the question, for it seems to derive from the human presumptuousness, many other manifestations of which are already familiar to

us. Nobody talks about the purpose of the life of animals (. . .) One can hardly be wrong in concluding that the idea of life having a purpose stands and falls with the religious system. (Freud 1930: 75)

Freud defends the same point of view in the *New Introductory Lectures*, where he defines a "Weltanschauung", a worldview, as

an intellectual construction which solves all the problems of our existence uniformly on the basis of one overriding hypothesis, which, accordingly, leaves no question unanswered and in which everything that interests us finds its fixed place. (Freud 1933: 158)

But then we learn that the world view of science

asserts that there are no sources of knowledge of the universe other than the intellectual working-over of carefully scrutinized observations—in other words, what we call research—and alongside of it no knowledge derived from revelation, intuition or divination. (Freud 1933: 159)

Philosophy "goes astray (. . .) accepting other sources of knowledge such as intuition" (Freud 1933: 160f). Not only is Freud hostile to questions such as the one about the meaning of life, he loathes inquiries that are discontinuous with science, which would include religious worldviews as well as, say, practical reasoning and constructivist approaches to morality. To use a term popularized by a group of theorists writing at the same time at the same place, such forms of discourse would be "meaningless".[11] ER is not an interesting subject because it provides an answer to a question (the question of the meaning of life) that, as far as Freud was concerned, cannot be raised intelligibly.

3.2 One can resist Freud's positivism in two ways: one may fault him for having too narrow a conception of objective inquiry, or else, regardless of how broad objective inquiry is in scope, object that Freud mischaracterizes its significance. While it is for reasons of the first sort that positivism is no longer a widespread view, Nietzsche rejects it for reasons of the second sort. He claims the success of the ascetic ideal is due to the absence of any successful alternative (EH III GM). While the most plausible competitor is science, Nietzsche devotes much of GM III to debunking this idea. "Science today," he says, "is not the opposite of the ascetic ideal but its most recent and noble manifestation" (GM III 23). Admiration of science does not commit him to Freudian positivism.

Nietzsche thinks science is committed to an "unconditional will to truth" (GM III 24). We can find two reasons why such an unconditional would be

[11] Freud concludes his discussion of worldviews by stating that "[a]ny of our fellow-men who is dissatisfied with this state of things, who calls for more than this for his momentary consolation, may look for it where he can find it. We shall not grudge it him, we cannot help him, but nor can we on his account think differently" (Freud 1933: 182). In light of these commitments, it is ironic that Freud himself was taken to task for being the embodiment of research that failed to satisfy standards of science.

ascetic, one in the *Gay Science*, and one in GM III. In GS 301 Nietzsche points out that we have created the world that concerns us. Already in GS 121 we read that life is no argument and that among the conditions of life may be the error. Within his program of de-deification and naturalization (GS 109) we can find a version of Kant's idea of the primacy of practical reason.[12] For Kant, practical reason has primacy over theoretical reason because of its insight into matters concerning freedom, God, and immortality. What Nietzsche would preserve of this view is that theoretical inquiry cannot offer answers to what matters in life. Thinking of science differently commits one to the ascetic ideal because it does not grant such primacy.

The second reason why the unconditional will to truth commits one to the ascetic ideal is that it is a will to non-perspectival truth, to truth as known from no particular perspective, a God's eye point of view. This is the topic of GM III 12, which famously claims that "[t]here is only a perspectival seeing, only a perspectival 'knowing' ". I follow Clark's (1990) reading of this passage in terms of a rejection of the view that knowledge is grounded in foundational beliefs all rational beings must accept. It is reasonable to equate non-perspectival knowledge with knowledge of things in themselves. Yet since in the *Gay Science* Nietzsche has abandoned the idea of things in themselves (GS 54), we can understand his endorsement of perspectivism as the epistemological counterpart to the metaphysical view expressed by the rejection of things in themselves. Absent such foundationalism, we reach the view that theoretical inquiry is influenced by the subject's cognitive constitution. If one loses sight of that, the will to truth becomes life-denying by ignoring the ways in which human concerns are integrated into one's cognitive perspective.

4. THE UNCONSCIOUS

4.1 Despite methodological similarities, Nietzsche-the-philosopher and Freud-the-scientist disagree about scope and significance of scientific inquiry. Freud's stance on science by itself would have prevented him from engaging with ER. However, Freud's model of the mind offers a reason to be skeptical about ER regardless of his positivism. Following Gardner (1993), an irrational person is one "who will in some loosely cognitive sense fail the test of self-confrontation: either he will be unable to provide an explanation-cum-justification of himself, or, in the course of attempting to do so, he will betray a failure of self-knowledge (p. 4)". Psychoanalysis explores what the mind must be like to be capable of the kinds of irrationality we observe, and applies this theory therapeutically where irrationality is strong enough to be an illness. Examples include hysteria, phobias,

[12] Cf. *Critique of Practical Reason,* Bk II, sect. 2.3.

anxiety, and obsessive-compulsive behavior, behavior Freud called neurotic.[13] Its erratic appearance notwithstanding, Freud held that such behavior can be causally explained in terms of mental states. He would explain neuroses from painful experiences embedded into the emotional lives of infants. Through repression, these experiences remain out of the mind's awareness before reemerging as patterns of behavior that appear inexplicable in the context of social interaction among adults.

Many philosophers have thought that human beings are not transparent to themselves, but it was Freud who offered systematic models of the mind based on the contrast between "conscious" and "unconscious". Yet "[h]ow are we to arrive at knowledge of the unconscious?" asks Freud (1915: 166), and continues: "It is of course only as something conscious that we know anything of it, after it has undergone transformation or translation into something conscious" (Freud 1915: 166). Still, Freud goes on to explain that the evidence for the existence of the unconscious is overwhelming, and pervasive in everyday life. Examples include parapraxes (Freudian slips); dreams; ideas that come to us but whose origins are mysterious; as well as latent memories and knowledge.

Until 1923, Freud used a topographical model of the mind, based on the opposition between conscious and unconscious parts conceived as regions of the mind. Only in the *Ego and the Id* does he propose a structural view, a tripartite model that related ego, id, and superego to one another and was supposed to offer a more suitable description of the phenomena.[14] The id is the repository of unconscious wishes and angers, the input of motivational material. Possessing the richest range of properties, the ego is charged with perception, believing, reality-testing, and may renounce instinctual satisfaction. Arising from the id through a process of identification (with individuals in the child's environment), the ego is the seat of anxiety as well as the source of repression. The superego, finally, emerges out of the child's ambivalent relationship to its caregivers. The child resolves the resulting anxiety resulting from both sexual and aggressive feelings through another act of identification, which gives rise to the superego. The superego monitors and directs the ego, by constituting a conscience and by holding up an ideal the ego feels compelled to match. Originally everything is unconscious; only under the influence of the external world does some of the mind's content become conscious.[15]

[13] I mostly follow Freud (1923), and Freud (1933). The approach to psychoanalysis in terms of what the mind must be like for observed forms of irrationality to become explicable draws on Gardner (1993).

[14] Freud (1923), section 1, is clear that there is an ambiguity in the term "conscious", which he uses in three different senses. This ambiguity was partly responsible for the reformulation of his theory in this work.

[15] See Wollheim (1971); Lear (2005); and Gardner (1993) for elaborations.

4.2 To see why this model delivers a reason to be skeptical of ER, note first that Freud's primary concern was with healing:

> Psychoanalytic therapy was created through and for the treatment of patients permanently unfitted for life, and its great triumph has been that by its measures a satisfactorily large number of these have been rendered permanently fit for existence. (Freud 1904: 263)[16]

Freud succeeds if the patient is restored to normal species-functioning. Concerned about criticism that psychoanalysis took too long, Freud responded that it had a pragmatically defined endpoint (cf. Freud 1937b). Yet a well-known formulation Freud uses to characterize psychoanalysis suggests a way of understanding psychoanalysis that renders its endeavors potentially open-ended. Lecture XXXI in the *New Introductory Lectures* ends by characterizing psychoanalysis in terms of "where id is there ego shall be" (Freud 1933: 80). This suggests an understanding of psychoanalysis as a tool for obtaining self-knowledge, not merely a device to render patients fit for existence. While healing processes might come to a pragmatically defined end, the creation of self-knowledge by turning "id into ego" may continue beyond that, and may be applied to persons who are not ill.

Acquiring self-knowledge, for Freud, generally involves others. After all, experiences were repressed *by the ego* because they were too painful. Although Freud analyzed himself, he questioned people's abilities to obtain self-knowledge by themselves. He even thought that "the direct utterances of the patient after he has been offered a construction afford very little evidence upon the question whether we have been right or wrong" (Freud 1937a: 263). This requires indirect forms of confirmation, highlighting the importance of the therapist. We see the therapist's role as yet more critical once we account for transference. Transference occurs when during the therapy memories reemerge and are repeated, with the analyst filling the role of somebody involved in them.[17]

So Freud's model of the mind by itself makes it plausible that self-knowledge will be hard to come by. Our inner lives are more complicated than introspection reveals. It is especially hard to acquire self-knowledge by oneself. But there is no point in endorsing ER unless one is transparent to oneself. Otherwise, one would confirm a life without knowing certain things about oneself that might make one choose differently. Since small doubts suffice to reject eternal repetition, it is hard to imagine how we could ever know enough to choose *this* life infinitely many times over all alternatives. Perhaps, one might say, psychoanalysis would

[16] For similar characterizations, cf. Freud (1937b: 250; 1937a: 257).

[17] Freud (1917: 444): "By that means the transference, which, whether affectionate or hostile, seemed in every case to constitute the greatest threat to the treatment, becomes its best tool, by whose help the most secret compartments of mental life can be opened." Cf. also Freud (1914).

put one into a position to know enough about oneself. But be that as it may, Nietzsche, at any rate, does not offer any such device.

4.3 Since Nietzsche had a similar model of the mind, these doubts about the possibility of endorsing ER apply to him. GM II offers an account of the development of the mind. GM II is concerned with the bad conscience, but as Risse (2001) argues, Nietzsche distinguishes between bad conscience as a feeling of guilt (the explanandum in GM II), and an early form of the bad conscience, an early form of the mind. The latter originates in people oppressed by intruders, the "pack of blond beasts of prey, a conqueror- and master-race", which "puts its terrible claws on a perhaps numerically vastly superior, but formless, still spreading population" (GM II 17). The oppressed are prevented from letting their instincts act against others, and the instincts are "enmity, cruelty, the lust for pursuit, for raid, for change, for destruction" (GM II 16). Nietzsche calls this inward-direction of instincts the *internalization* of man. The effects of this development are immense. It is only through the rise of the older form of the bad conscience that we can understand "contradictory notions" (GM II 18) such as selflessness, and self-sacrifice as ideals. Only after the oppression of instincts is there a sufficiently rich inner life to allow for ideas such as the "un-egoistic". From then on, a person is a plurality owing to the presence of competing instincts.

Nietzsche also has views about how the mind, once emerged, functions, which he develops in many remarks across his works. Three views stand out: first, the workings of the mind are shaped by instincts; second, we are unconscious of much in our minds; and third, consciousness is a social phenomenon in a sense to be explained below. As far as the first point is concerned, Nietzsche agrees with Freud, but I will not elaborate on that issue.[18] The second theme is present across Nietzsche's writings. In D 116, we find Nietzsche emphasizing (in a section titled "The unknown world of the 'subject' ") that "what people find hard to grasp is their ignorance about themselves". A little later, in D 129, Nietzsche explains in what sense we remain unknown to ourselves:

> What takes effect is the habit of our play of forces, or a little impetus from a person whom we fear or honor or love, or the convenience that prefers to do what is at hand, or the excitement of fantasy, brought about by the next available smallest event at the crucial moment, what takes effect is the physical, which appears entirely unexpectedly, what takes effect is the mood (. . .): in brief, what takes effect is motives that we partly do not know at all and partly very badly and that we can never ponder against each other in advance.

[18] For Freud, see Freud (1920; 1923: ch. 4; 1933: XXXII). For Nietzsche, see D 119; GS 333; BGE 3, BGE 6, BGE 12, BGE 19, BGE 117. On Nietzsche, see also Parkes (1994); Thiele (1990: 51 ff); and Richardson (1996), ch. 1.4, for the drive constitution of the human mind. While Nietzsche rarely discusses sexuality, he points out that "the degree and type of the sexuality of a man extend all the way to the ultimate peak of his spirit" (BGE 75).

BGE 17 takes up on this theme that our motivations remain largely beyond our control. BGE 17 is, as best as we can tell, the origin of the Freudian "id":[19]

So far as the superstitions of the logicians are concerned, I will never tire of emphasizing over and over again a small brief fact which these superstitious types are unhappy to concede—namely, that a thought comes when "it" wants to and not when "I" want it, so that it's a falsification of the fact to say that the subject "I" is the condition of the predicate "think". It thinks: but that this "it" is precisely that old, celebrated "I" is, to put it mildly, only an assumption, an assertion, in no way an "immediate certainty". After all, we've already done too much with this "it thinks": this "it" already contains an interpretation of the event and is not part of the process itself.

Let me explain the third theme, that consciousness is a social phenomenon. Nietzsche explains the phenomenon well, so I quote GS 354 at some length:[20]

[C]onsciousness has developed only under the pressure of the need for communication; [. . .] from the start it was needed and useful only between human beings (particularly between those who commanded and those who obeyed). [. . .] Consciousness is really only a net of communication between human beings; it is only as such that it had to develop; a solitary human being who lived like a beast of prey would not have needed it. [. . .] Man, like every living being, thinks continually without knowing it; the thinking that rises to consciousness is only the smallest part of all this—the most superficial and worst part—for only this conscious thinking takes the form of words, which is to say signs of communication, and this fact uncovers the origin of consciousness. [. . .] The human being inventing signs is at the same time the human being who becomes ever more keenly conscious of himself. It was only as a social animal that man acquired self-consciousness—which he is still in the process of doing, more and more. [. . .] My idea is, as you see, that consciousness does not really belong to man's individual existence but rather to his social or herd nature; that, as follows from this, it has developed subtlety only insofar as this is required by social or herd utility. Consequently, given the best will in the world to understand ourselves as individually as possible, "to know ourselves", each of us will always succeed in becoming conscious only of what is not individual but "average". Our thoughts themselves are continually governed by the character of consciousness—by the "genius of the species" that commands it—and translated back into the perspective of the herd. Fundamentally, all our actions are altogether incomparably personal, unique, and infinitely individual; there is no doubt of that. But as soon as we translate them into consciousness *they no longer seem to be*. [. . .] [W]hatever becomes conscious *becomes* by the same token shallow, thin, relatively stupid, general, sign, herd signal; all becoming conscious involves a great and thorough corruption, falsification, reduction to superficialities, and generalization.

So while much of our mind is inaccessible, what is accessible at all is the mind's "public" area: those regions (thinking topographically) or aspects (thinking

[19] Gasser (1997: 107–18), offers an account of a view common among Freudians, that the term "id" "had been extensively employed by Nietzsche" (Jones 1953: vol. 3, 280). While Nietzsche does not "extensively" employ anything that might correspond to Freud's "id", the origin of this usage can be traced to BGE 17.

[20] On consciousness, see also GS 11 and BGE 3.

systematically) that have emerged to facilitate communication and that are harmonized with other minds. What is individual is unconscious. While Nietzsche does not mention mechanisms such as repression, the sense in which a person's individuality is unconscious is not what Freud called "descriptive" and hence accessible to reflection; it is unconscious in the Freudian sense of "dynamic", and thus even upon trying not easily, if at all, recoverable.

4.4 Recall Nietzsche's encouragement to "become what you are" (cf. GS 270, GS 335). In light of his and Freud's conceptions of the mind, we see how similar this is to Freud's "where id is, ego shall be". But on Nietzsche's view, parallel to what we found about Freud above, realizing this admonition, or for that matter endorsing ER, is difficult. Those parts of the mind that matter for who we are *individually* are closed to us. Consider a person confronted with the question of whether she would endorse an infinite repetition of her life. What she must assess is whether she would be willing to return as the same person again and again. But lacking self-knowledge means lacking insight into relevant aspects about how one's life will unfold. How one's life unfolds depends on circumstances beyond one's control, but also on how one would react to such circumstances—and it is knowledge of this sort that an agent does not have. So confronted with the demon's question of whether she would want to live precisely this life again and again, she would have to say: "I do not fully know who I am, and thus do not understand what you refer to when talking about 'this' life, so how can I answer you?"

To dispel this worry, one might suggest that, while the agent may face uncertainty vis-à-vis her life, the scope of possibilities of what lies hidden in one's unconscious is limited. Perhaps the agent could endorse the whole range of possible lives so delineated. Yet it is hard to see how there could be a wholehearted endorsement of a scope of possible lives (at any rate if, as we can safely presume, they differ in ways the agent herself would consider significant). The kind of enthusiasm required for the agent to respond to the demon by considering the thought of ER "divine" requires that it be well understood what "this" life is. Nietzsche does not offer a theory of how what is hidden from us may be recoverable. So, on his view, it is *even harder* to see than on Freud's how anybody would be able to endorse an infinite repetition of his life. There will always be at least sparks of doubt kindled by these insights into how the mind works, and such sparks are enough to discourage one from responding affirmatively to the demon's question. The demon, after all, does not simply ask whether the agent would be willing to live her life again; he asks whether she would want to relive it infinitely many times, to make sure there are not even such sparks of doubt, sparks whose significance would be magnified through this appeal.[21]

[21] One may say Nietzsche has a response to these queries. After all, whereas Freud was concerned with working-through, Nietzsche seemed eager to seal off, to forget. But for one thing one may

A different way of dispelling the worry is by saying that affirmability does not turn on self-knowledge. The demon, one may say, tests one's willingness to affirm one's life by promoting a response that is not reflective to such an extent that complete self-knowledge would be required. What is measured is plain enthusiasm. Two initial responses are, first, that sheer enthusiasm cannot be a proper way to think about ER because that would make it too easy to endorse; if the Epsilons in Huxley's *Brave New World* find themselves in a position to affirm their lives, something has gone wrong. Second, the demon speaks to the agent in her "loneliest loneliness", which, like Butler's "cool hour" (see his *Fifteen Sermons*, Sermon IV) seems designed as a moment for reflection. A third, stronger, point is that any such response will leave the agent without a serious answer to a second-guessing challenge that points out to her that she does not understand well enough what she said she would be willing to repeat over and over again. No matter how precisely one spells out what "affirmation" amounts to, this would be an unsatisfactory feature of it. If infinite repetition is at stake, knowing something well enough is not good enough. For now, I regard this move as unsatisfactory, but will return to it in the concluding section.[22]

5. GUILT

5.1 So we have found a reason why Freud would reject ER that also applies to Nietzsche. There is one more such reason, which turns on their respective accounts of guilt. Freud offers both a phylogenetic and an ontogenetic account of guilt, that is, an account explaining the emergence of guilt within the human species as well as one concerning individuals.[23] Let us first glance at the phylogenetic account in Freud (1913). Following Darwin, Freud thought that, originally, humans lived together in hordes ruled by a strong male. This male had sexual access to all females and drove away other males, including his

wonder whether this can be done successfully (cf. D 126); and for another thing, on his account in GS 354, what will be sealed off in this way is what makes the individual into that actual individual. At any rate, the issue of praise for forgetting versus praise for memory is orthogonal to issues of self-knowledge. One may also say that one hardly ever knows anything entirely, in the sense that there will always be properties of an object that will remain hidden; but nevertheless, we succeed in integrating objects, even persons, into our lives even without such complete knowledge. So why should it be problematic that we do not fully understand what "this" life is that we are asked to affirm? But this question now has a ready answer: because the question is whether we would want to relive this particular life over and over again. I explained above why, for such purposes, lack of knowledge is problematic, whatever else is true for other cases of lack of knowledge.

[22] Nehamas (1985) understands an endorsement of ER to require an endorsement of an eternal repetition of everything, not just of an endorsement of all aspects of one's personality. On that understanding, of course, endorsing ER becomes difficult for yet a different reason. This is, however, a reading of the text that I do not endorse, though here is not the time to explore the reasons.

[23] Freud (1930) introduces the phylogenetic account in sect. IV, and the ontogenetic one in sect. VII. See in particular p. 131. See also Freud (1913: 156–7).

offspring once of age. Those expelled males banded together to slay the father. Afterwards, their love for the father resurfacing, they set him up as a totem. As Freud put it, "the affection which had all this time been pushed under was bound to make itself felt. It did so in the form of remorse. A sense of guilt made its appearance" (Freud 1913: 143). They also instituted prohibitions on killing and incest. These developments (which occurred often enough to gain importance) mark the beginning of religion and morality.

There are two versions of the ontogenetic account, one in *Civilization and Its Discontents*, the second in all other relevant writings. According to the former, guilt emerges from emotional ambivalence vis-à-vis parents. Children love their caregivers, but develop hostility towards them as authorities who frustrate their desires. Yet they cannot act on such anger for fear of losing the parental love. The child resolves this ambivalence by incorporating the parents into her personality and investing the newly-formed part of her personality with all hostility that has arisen. (The result is the superego.) The conscience arises from a redirection of aggressiveness from the parents and towards the child herself:

> A considerable amount of aggressiveness must be developed in the child against the authority which prevents him from having his first, but none the less his most important, satisfactions, whatever the kind of instinctual deprivation that is demanded of him may be; but he is obliged to renounce the satisfaction of this revengeful aggressiveness. He finds his way out of this economically difficult situation with the help of familiar mechanisms. By means of identification he takes the un-attackable authority into himself. The authority now turns into his superego and enters into possession of all the aggressiveness which a child would have liked to exercise against it. (Freud 1930: 129)

Deigh (1996) calls this the *Nietzschean* account of the development of the superego. The other account (e.g., Freud 1923; 1933) differs by its emphasis on sexual instinct, rather than subjection to authority. Infants form strong attachments to parents. Later, they form an erotic attachment to one parent while identifying with the other. The child comes to regard the parent with whom she identifies as a rival for the affection of the other and develops aggression towards the former. This situation of ambivalence is the Oedipus complex. The child resolves this ambivalence by abandoning his erotic attachment and identifying with the parent towards whom the attraction is felt, or, more often, by intensifying his identification with the other.[24] Deigh calls this the *Standard Account*.

Various questions arise. First, we must wonder about the coherence between the phylogenetic and the ontogenetic account. In the former, some version of guilt is already present: after all, the brothers feel remorse. Yet in the latter, the

[24] Freud (1913: 24–5; 1933: 64). Lear (2005: 180–3) points out that Freud often fails to discuss the Oedipus complex at the right level of generality, so that it seems as if the specific attachments a child has to the parent of the one and the parent of the other sex appear to be more important than they really are. What matters is general ambivalence that is partly sexual in nature.

sense of guilt (explained in terms of the tension between ego and superego[25]) is itself explained alongside morality. Moreover, we must make sense of the fact that Freud offers two versions of the ontogenetic account, which we cannot do by thinking he changed his mind. For the Standard Account appears both before and after the Nietzschean Account.[26]

But these difficulties need not trouble us. What matters is that "the price we pay for our advance in civilization is a loss of happiness through the heightening of the sense of guilt" (Freud 1930: 134). For this sense of guilt, on any account that could stem from Freud, "is set going as soon as men are faced with the task of living together" (Freud 1930: 132). Conscience is a device by which civilization turns to its advantage the aggressiveness of our native endowment, whose unimpeded realization would create an environment too hostile for civilization. Conscience is like "an agency within [man] to watch over [the aggressive instinct], like a garrison in a conquered city" (Freud 1930: 124).[27] Reason can do no more about the conscience than the city about its garrison. This gives us another reason why Freud would reject ER: the bad conscience is part of what it is to live in civilization, and so is guilt. The presence of guilt does not commit us to an attitude Frida Kahlo took at the end of her life, when she wrote in her diary that she hoped the end would be joyful and that she would never return. But this account does point to elements of our emotional lives we will not want to have repeated but cannot be without.

5.2 Back to Nietzsche. Nietzsche's account of guilt in GM contains three components:[28] To begin with, there is an early form of the bad conscience, which is an early form of consciousness. The second element is indebtedness to ancestors and gods. This indebtedness arises in rituals of worship, but is transformed into a domineering feeling that turns into guilt. Nietzsche characterizes this transformation as a "pushing-back" (GM II 21) of indebtedness into the original form of the bad conscience. This pushing-back is brought about by Christianity, the third component. I have already discussed the first element. Let us consider the second element. Having debts is a juridical relationship, and the emotional and moral connotations of *guilt* do not pertain to this original relationship. One variant is the debt of offspring towards ancestors for their contributions to the tribe (GM II 19). The debts grow the more the tribe succeeds; eventually, ancestors transfigure into gods. Yet moralization occurs only through the "pushing-back of

[25] Cf. Freud (1923: 50f; 1933: 78; 1930: 135). [26] See Deigh (1996) for helpful discussion.

[27] Cf. Nietzsche, BGE 188: "Every morality is—in contrast to *laisser aller*—a part of tyranny against 'nature,' also against 'reason:' that is, however, not yet an objection to it".

[28] This is developed in detail in Risse (2001). The similarities between Nietzsche and Freud on the subject of guilt have been recognized at least since Jones (1953: vol. 3, 283–4), and the core of these similarities is the prominence of a process of internalization. However, Jones is also aware that Nietzsche's account is an exclusively phylogenetic one.

those notions into the conscience, or more specifically, through the involvement of the *bad* conscience with the concept of God'' (GM II 21).

To explain this ''pushing back'', we need to introduce Christianity. As Nietzsche puts it in GM III 20, ''only in the hands of the priest, this true artist of the feeling of guilt, did it take on form—oh what a form! 'Sin'—for thus reads the priestly reinterpretation of the animal's 'bad conscience' cruelly turned backwards''. The priest invents an ethical world order (A 26). Many of man's natural instincts come to be seen as dispositions to violate this order, that is, as *sins* (GM III 20).[29] The ''pushing-back'' is a psychological consequence of accepting Christianity. Its endorsement generates a new sentiment, *guilt*, which is so strong that it, in turn, generates a new kind of ''inner life''. The original indebtedness turns into a deep sense of failure with respect to what one is first and foremost, namely, God's creature.

In some ways, Nietzsche's account differs remarkably from Freud's. First, what Nietzsche explains through internalization is not guilt, but an early form of the mind. Guilt enters only through Christianity. Second, guilt, for Nietzsche, may not be present in civilization. He makes the point by referring to the Greeks (GM II 24), who used gods differently (blaming them for mishaps) and thus were not bothered by guilt. Third, Nietzsche fails to offer an ontogenetic account to accompany his phylogenetic account. He does not say, for instance, whether guilt is passed on genetically, or through adaptation. Fourth, guilt is not tied to the emergence of morality in the same way in which it is for Freud. It is through *ressentiment* that the moral standpoint arises. For these reasons, calling the account in *Civilization and its Discontents* ''Nietzschean'' is misleading.

5.3 So while for Freud individuals cannot overcome guilt, Nietzsche disagrees, and thus seems to escape from that difficulty for endorsing ER. Yet appearances are deceiving. For the moral psychology in GM also includes *ressentiment.* It is not often noticed that one challenge GM poses is to offer an account of *ressentiment* that locates its development within the philosophical anthropology in GM II. Questions about the origin of *ressentiment* may seem unmotivated since its occurrence does not appear puzzling: slaves are oppressed, and what more is required to explain their resentment? Yet guided by our own emotional lives, *we* may or may not find the emergence of *ressentiment* puzzling, but Nietzsche *must* account for it. For the goal of the *Genealogy* is to develop the kind of ''animal psychology'' (GM III 20) that explains why our emotions are what they are. Nietzsche would rest his case too early if he could not ground *ressentiment* within his anthropology.

[29] Cf. A 26: ''What does 'moral world order' mean? That there is a will of God, once and for all, as to what man is to do and what he is not to do; that the value of a people, of an individual, is to be measured according to how much or how little the will of God is obeyed; that the will of God manifests itself in the destinies of a people, of an individual, as the ruling factor, that is to say, as punishing and rewarding according to the degree of obedience.''

Since Risse (2003) offers an account of *ressentiment*, I only stress a few points. *Ressentiment* is a state of mind that arises when the mind becomes self-conscious under circumstances where deep-rooted anger and resentment are already present. Anger and resentment arise because the mind evolves in response to the oppression of aggressive instincts. For the mind to become conscious of itself means for it to become able to refer to and reflect upon itself, perhaps by forming representations of itself or beliefs, emotions, or other entities "inside" or constitutive of it, perhaps in other, non-representational ways. Once the mind is capable of referring to itself, anger and resentment find new targets inside of it, becoming directed against the mind itself. Nietzsche may say of *ressentiment* what he says of guilt: that it "fixes itself firmly, eats into him, spreads out, and grows like a polyp in every breadth and depth" (GM II 21). It is through *ressentiment* (as much as through guilt) that "the human soul became *deep*" (GM I 6)—a remark that has a peculiar literalness to it.

The most significant statement about the origins of *ressentiment* appears in GM III:

> It is here alone, according to my surmise, that one finds the true physiological causality of *ressentiment,* of revenge, and of their relatives—that is, in a *longing for anesthetization of pain through affect*:—this causality has been commonly sought, very mistakenly to me, in the defensive counterblow, a mere reactive protective measure, a "reflex movement" in the case of some sudden harm and endangerment, of the kind that a frog without a head still carries out in order to get rid of a corrosive acid. But the difference is fundamental: in the one case, one wishes to prevent further damage, in the other case, one wishes, by mean of a more vehement emotion of any kind, to anesthesize a tormenting, secret pain that is becoming unbearable and, at least for the moment, to put it out of consciousness. (GM III 15) [my emphasis]

Ressentiment, then, arises as a physiological reaction against the pain due to the process of civilization. *Ressentiment* numbs the pain that arises through forced socialization. Yet since aggressive instincts characterize human beings as they are by nature, anger and resentment are typical phenomena accompanying civilization, and it is bound to be the exception that an individual has overcome or avoided the tormented state thus produced.

For Nietzsche *ressentiment*, rather than guilt, is the price of civilization. *Ressentiment,* to use a formulation Freud employed for guilt, is "set going as soon as men are faced with the task of living together" (Freud 1930: 132). But the person of guilt and *ressentiment* must be understood as opposed to the one capable of endorsing ER. So we have encountered another reason why agents would have a hard time endorsing ER. For Freud, this reason turns on guilt, whereas for Nietzsche, it turns on *ressentiment,* but both notions share their centrality to civilization on the respective view.

6. "THE GREATEST WEIGHT"

We have seen that Freud rejects—or would have rejected, had he bothered—ER for at least three reasons, reasons it is useful to explore because Nietzsche and Freud share methodological commitments. The first one turns on basic epistemological commitments. Whereas Freud and Nietzsche differ substantially with regard to these commitments, the two other reasons draw on views they share. We saw that endorsing ER will require complete transparency (to avoid nagging doubts about what one could be), as well as an overcoming of guilt and *ressentiment* (to overcome what otherwise would constitute a reason against repeating one's life as it is). Since, however, neither is *impossible* to achieve, the naturalist reading, and the accompanying need to spell out some of Nietzsche's views in more detail than he did, does not render his emphasis on ER unintelligible. No inconsistency lurks, and thus the tensions we have found in this study may well be resolvable. Let me conclude by considering various strategies for doing so.

One possible resolution begins by pointing out that, unlike Freud, Nietzsche was not interested in assisting ordinary people, but in a selected few. What this study illuminates is why it will be *extraordinarily* exceptional for anybody to be among those selected few, and Nietzsche's views on the nature of the mind and his philosophical anthropology show why this is so. Yet the problem with this resolution is that part of it is not very Nietzschean. Sure enough, the higher men are supposed to overcome guilt and *ressentiment*, and would be rather exceptional for it. But while Nietzsche often emphasizes that we lack self-knowledge, he never ties the attainment of complete self-knowledge to his ideal of the higher men, or to any other recognizable ideal in his works.[30]

Another resolution insists that this sort of inquiry about ER is misguided. What is required is precisely an unreflective endorsement of ER, or at any rate none that loses itself in analytical subtleties. Clark (1990: 268), for instance,

[30] There are, however, in Nietzsche's late works various references that emphasize the importance of what is "real". In EH, Nietzsche complains that "one has deprived reality of its value, its meaning, its truthfulness, to precisely the extent to which one has mendaciously invented an ideal world. (. . .) The *lie* of the ideal has so far been the curse on reality" (*Preface* 2), and similar formulations in which Nietzsche attaches value to the real apparently because it is real can be found in TI *Skirmishes*, 32 ("How much greater is the worth of the real man, compared with any merely desired, dreamed-up, foully fabricated man? with any ideal man? And it is only the ideal man who offends the philosopher's taste."); A 15 (where Nietzsche complains that Christianity "de-values" the real world); as well as A 9 and EH, *Fate*, 7. If what is real is so important, one might say, then so is self-knowledge, as knowledge of one part of what is real. Strikingly, however, these passages emphasizing the importance of what is real never stand in any connection with self-knowledge, and those passages in which lack of self-knowledge is described do not contain a clear indication that there is anything deeply problematic about it.

argues that one should think of the demon's question along the lines of a question one might ask of one's spouse: "Would you marry me all over again?" In this case affirmation would amount to a spontaneous and resounding "yes". Reflection would not make one's answer more solid or credible; on the contrary, reflection is not what it takes to assess whether one loves somebody, and a request to be granted time for reflection would invalidate any positive answer that might emerge later. But while it might well be in the nature of love that its presence can be tested in such ways, genuine enthusiasm for one's life cannot be. There would be nothing about a request to be granted time for reflection that would undermine any positive answer that might emerge through such reflection; on the contrary, given the complex nature of what a person's life is, reflection is what it takes to come to an assessment of it, and an affirmation of one's life that emerges through reflection would indeed be more solid and credible for it. At any rate, a life affirmation test that actually *depends on* insisting that reflection be suppressed would be rather unpersuasive. This sort of test seems suited to ascertain a sunny disposition rather than a more deeply-seated attitude towards one's life.

A third strategy for resolving the tensions would be to reflect on what precisely it is that is being affirmed. We have assumed that what is affirmed is a "life" in the sense of "being a certain person who leads a certain life". And if indeed the test asks whether one can affirm the kind of person one is, one is easily led to wonder whether one understands oneself well enough to predict how one would react in certain situations. One may well have desires to be the kind of person who reacts in admirable ways under extreme conditions (to be courageous, remain calm under severe pressure, be loyal in response to overwhelming enticement, etc.), and to the extent that one may do so, one may wish one were a different person if one does not have the disposition to react in such ways; and if, in turn, one does not know such things about oneself one might doubt one's ability to affirm one's life. But what if one is merely asked to affirm the set of events of which, up to *now,* one's life has been composed? On such an approach one would not worry about dispositions that have not revealed themselves *because* they have not revealed themselves.

Yet the price to pay for avoiding the concerns of this study in this way is that one would affirm one's life only in a highly qualified sense, and would do so rather consciously unless one combined this view with the second attempted resolution (and would inherit its problems as well). It might just so happen that one has lived to find oneself in a happy set of circumstances and thus would affirm one's good luck more than one's character by answering the demon positively. Moreover, doing so would actually set aside the deeper question of whether one would like to return *as this person*, replacing it instead with a question that is easier to answer. But then the enthusiasm one is obtaining through this test for one's life could not be deep.

Whereas the first attempted resolution was wanting because it did not seem to be Nietzsche's proposal, the second and the third were wanting because they

compel us to think of Nietzsche's own enthusiasm for his allegedly greatest insight as overdrawn. In response to these difficulties, a fourth solution tries to capture the paradoxical thought that ER asks for an affirmation of something that cannot be affirmed. On this view, one should endorse ER although Nietzsche's other views do indeed provide good reasons why one should not endorse it. Affirmation would be required without affirmability, and it is this that makes an endorsement of ER "the greatest weight". While this proposal could arguably avoid the drawbacks of the first three, it returns us to the difficulties we encountered above when discussing Nietzsche's views on the ability of attaining self-knowledge, and we are not closer to resolving them. Of course, above we tried to make an endorsement of ER sensible, and on this view—*affirmation without affirmability*—that attempt would be abandoned. Affirmation, then, would be a matter of irrational enthusiasm for one's life, an enthusiasm not open merely to Nietzsche's higher men but also to Huxley's Epsilons. That by itself would not rule out this resolution because the higher men can also be characterized in ways other than their ability to affirm ER, although, then, among other things, they would be of the kind of character that could engage in this affirmation without affirmability. Yet there would remain something to be desired about an account of ER that requires a mindset for agents in which they both affirm ER and know they have insufficient reason to do so.

None of these four approaches, then, is entirely satisfactory. We find ourselves without a resolution that can be considered Nietzschean in spirit, would not turn Nietzsche's view that the Eternal Recurrence was his greatest insight into an exaggerated stance, and would still be an overall plausible view. The demon in GS 341 is a rather vexing character, not just practically, but also theoretically.

REFERENCES

Works of Nietzsche are from the *Kritische Studienausgabe*, edited by Giorgio Colli and Mazzino Montinari, dtv/de Gruyter, second edition 1988. I use the usual abbreviations for the works in English translation, i.e., GM for *On the Genealogy of Morality*, TI of *Twilight of Idols*, D for *Dawn*, HAH for *Human All Too Human*, WS for *The Wanderer and his Shadow* (part of *Human, All too Human* II) GS for *Gay Science*, A for *Antichrist,* EH for *Ecce Homo*, Z for Zarathustra, and BGE for *Beyond Good and Evil.*

Works of Freud are from *The Standard Edition of the Complete Psychological Works of Sigmund Freud,* (CPW), trans. James Strachey in collaboration with Anna Freud (London: Hogarth Press, 1953). I use the following abbreviations for Freud:

(1904) "On Psychotherapy," pub. 1904, CPW, vol. 7, 257–68

(1913) *Totem and Taboo*, pub. 1955, CPW vol. 13, 7–162

(1914) "Further Recommendations in the Technique of Psychoanalysis: Recollection, Repetition, and Working Through," pub. 1958, vol. 12, 145–56

(1915) "The Unconscious," pub. 1957, vol. 14, 159–215

(1917), *Introductory Lectures on Psychoanalysis*, vols. 15 and 16, pub. 1961 (vol. 15) and 1963 (vol. 16)
(1920) *Beyond the Pleasure Principle*, pub. 1955, vol. 18, 3–64
(1923) *The Ego and the Id,* pub. 1961, vol.19, 3–66
(1927) *The Future of an Illusion*, vol. 21, pub. 1961, vii–56
(1930) *Civilization and Its Discontents*, vol. 21, pub. 1961, 57–145
(1933) *New Introductory Lectures on Psychoanalysis*, vol. 22, pub. 1964, 3–182
(1937a) "Constructions in Analysis," pub 1964, vol. 23, 255–69
(1937b) "Analysis Terminable and Interminable," pub 1964, vol. 23, 209–53.

OTHER WORKS CITED

Assoun, Paul-Laurent (2000). *Freud and Nietzsche* (London: Continuum).
Chapelle, Daniel (1999). 'Nietzsche and Psychoanalysis: From Eternal Return to Compulsive Repetition and Beyond', in Jacob Golomb, Weaver Santaniello, and Ronald Lehrer (eds.), *Nietzsche and Depth Psychology* (Albany: State University of New York Press), 37–51.
Chapman, A. H., and Mirian Chapman-Santana (1995). 'The Influence of Nietzsche on Freud's Ideas', *British Journal of Psychiatry* 166 (2), 251–3.
Clark, Maudemarie (1990). *Nietzsche on Truth and Philosophy* (Cambridge: Cambridge University Press).
Danto, Arthur (1965). *Nietzsche as Philosopher* (New York: Macmillan).
Deigh, John (1996). 'Remarks on Some Difficulties in Freud's Theory of Moral Development', *The Sources of Moral Agency* (Cambridge: Cambridge University Press), 65–93.
Gardner, Sebastian (1993). *Irrationality and the Philosophy of Psychoanalysis* (Cambridge: Cambridge University Press).
Gasser, Reinhard (1997). *Nietzsche und Freud* (Berlin: de Gruyter).
Gödde, Günther (1999). *Traditionslinien des "Unbewussten." Schopenhauer—Nietzsche—Freud* (Tübingen: Edition Diskord).
Greer, Scott (2002). 'Freud's "Bad Conscience": The case of Nietzsche's Genealogy', *Journal of History of the Behavioral Sciences* 38 (3), 303–15.
Hatab, Lawrence (2005). *Nietzsche's Life Sentence. Coming to Terms with Eternal Recurrence* (New York: Routledge).
Jones, Ernest (1953). *The Life and Work of Sigmund Freud*, 3 vols. (New York: Basic Books).
Lear, Jonathan (2005). *Freud* (New York: Routledge).
Lehrer, Ronald (1995). *Nietzsche's Presence in Freud's Life and Thought* (Albany: State University of New York Press).
Leiter, Brian (2002). *Nietzsche on Morality* (London: Routledge).
Löwith, Karl (1978). *Nietzsches Philosophie der Ewigen Wiederkehr des Gleichen*. 3rd edn. (Hamburg: Meiner).
Magnus, Bernd (1978). *Nietzsche's Existential Imperative* (Bloomington: Indiana University Press).
Moore, Gregory (2002). *Nietzsche, Biology, and Metaphor* (Cambridge: Cambridge University Press).

Nehamas, Alexander (1985). *Nietzsche—Life as Literature* (Cambridge, Mass.: Harvard University Press).
Richardson, John (2004). *Nietzsche's New Darwinism* (Oxford: Oxford University Press).
Risse, Mathias (2001). 'The Second Treatise in *On the Genealogy of Morality*: Nietzsche on the Origin of the Bad Conscience', *The European Journal of Philosophy,* 9, 55–81.
——(2003). 'Origins of *Ressentiment* and Sources of Normativity', *Nietzsche Studien* 32, 142–70.
Scavio, M. J., A. Cooper, and Clift P. Scavio (1993). 'Freud's Devaluation of Nietzsche', *Psychohistory Review* 21 (3) 295–318.
Stambaugh, Joan (1972) *Nietzsche's Thought of Eternal Return* (Baltimore: Johns Hopkins University Press).
Wollheim, Richard (1971). *Sigmund Freud* (Cambridge: Cambridge University Press).

This page intentionally left blank

12

Nietzsche on the Will: An Analysis of BGE 19

Maudemarie Clark and David Dudrick

Given the apparent importance of the will in Nietzsche's philosophy, one would like to know what he thinks the will is. Yet his single most important passage on the will, *Beyond Good and Evil* 19, although well known and frequently quoted, has received little illuminating analysis. This is actually not surprising, given how difficult it is to make sense of what Nietzsche says in it. Particularly puzzling is its ending, of which, to our knowledge, no one has even attempted an account. But its beginning already presents a puzzle for careful readers: why does Nietzsche's suggest that he intends to provide an alternative to traditional philosophical accounts of the will, an analysis of what it is to will or have a will, when he apparently offers instead a phenomenology of willing, an account of the *experience* of willing?

We finally have an illuminating interpretation of BGE 19 in Brian Leiter's contribution to this volume, which also offers a plausible answer to this question.[1] According to Leiter, Nietzsche denies that there is any such thing as a will. He offers a phenomenology of willing to show that the experience that leads us 'to conceive of ourselves as exercising our will' fails to '*track* an actual causal relationship' and that it 'systematically misleads us as to the causation of our actions' (Leiter 2009: 122). Leiter thus takes BGE 19 to offer a deflationary account of the will. Willing exists only in the sense that we have an experience of willing. Willing is therefore epiphenomenal: it belongs to consciousness but has no effect on behavior. We'll call this the naturalistic interpretation of BGE 19 because it supports Leiter's naturalistic take on Nietzsche, according to which 'human actions and beliefs (including beliefs about morality)' are to be explained causally, in terms of the person's 'fixed psycho-physical constitution'.[2] According to this naturalistic picture, the will is an unnecessary posit, just another item to be thrown on the junk-pile of outlived metaphysical fantasies.

[1] Leiter, Brian (2009). 'Nietzsche's Theory of the Will', (this volume). The translations of BGE are our own, but we have consulted and learned from those of Hollingdale, Kaufmann, and Norman.

[2] Leiter, Brian (2002). *Routledge Philosophy Guidebook to Nietzsche on Morality*. (London and New York: Routledge), 8.

The naturalistic interpretation offers a plausible way to read BGE 19; indeed we think it is the way in which Nietzsche's writing sets us up to read it. But we do not think it makes the best sense of this passage. BGE 19 thus exemplifies a pattern we find throughout BGE: Nietzsche writes in a way that almost inevitably seduces readers into a naturalistic interpretation of his claims, but he also embeds certain features in the passage that cannot be made good sense of in naturalistic terms. The point, we claim, is to lead careful readers to a philosophically more sophisticated interpretation of Nietzsche's claims in normative rather than purely naturalistic—strictly causal—terms.

Our strategy for demonstrating the plausibility of the normative reading of BGE 19 is to devote ourselves to bringing out the inadequacies of the naturalistic interpretation of the passage. Because Leiter has provided the fullest development of and support for the latter, we devote much of the present paper to arguing against his interpretation, bringing out our alternative normative interpretation of Nietzsche's will in contrast to his. Our paper has four sections. We disagree with Leiter's account of Nietzsche's phenomenology of willing in Section 1, and with his account of its function in Section 2. Its function, we argue, is not to debunk the will, showing that our experience of willing fails to '*track* an actual causal relationship', but to provide a basis for a theory of what willing actually is. Willing, as we present Nietzsche's account, is possible only for a being that operates according to values. Once we recognize this, it becomes implausible that the passage aims to show that willing is epiphenomenal. In Section 3, we concede that Nietzsche *is* rejecting something about the traditional conception of the will in BGE 19, but explain what that is in a way that leaves the causality of the will intact. In Section 4 we address questions concerning the plausibility of the theory of the will we have attributed to Nietzsche on the basis of BGE 19 and 21.

Our analysis of these passages supports the claim, for which we have argued elsewhere (Clark and Dudrick 2005), that although Nietzsche is a naturalist in an important sense, and certainly rejects all forms of supernaturalism, there is an important sense in which he is not a naturalist, for he holds that human thought and action can be understood only from a perspective constituted by norms that have no role to play in our understanding of the natural world. It also provides evidence for the related claim that Nietzsche has greater sympathy with traditional metaphysical notions than has been recognized. In particular, BGE 19 aims to rehabilitate the traditional notion of the will in the face of the tendency of naturalism to simply dismiss it. Our forthcoming book argues that earlier sections of BGE undertake this kind of rehabilitative project for the notion of the soul. Though we can't pursue the point here, we take BGE 6, 9, and 12 to show that Nietzsche takes *who a person is* to be constituted by the order of her drives, an order that is not merely causal but normative: a ranking according to authority and not simply strength. In accord with the traditional conception, Nietzsche takes the will to be the soul, insofar as the soul is involved

in bringing about action. We'll argue here that BGE 19 presents a person's willing as constituted by certain activities of this normative order. The upshot is that Nietzsche's attitude towards traditional metaphysics is not one of outright rejection. He certainly rejects anything in such notions that is incompatible with what natural science tells us, or that can only be defended on a priori grounds. But he does not dismiss such notions as nonsense. The soul, for instance, is one of the 'most ancient and venerable of hypotheses', and Nietzsche does not want to 'forgo' this hypothesis by 'get[ting] rid of the soul itself', but to offer one of the possible 'new forms and refinements of the soul-hypothesis' (BGE 12). This refinement would make it possible for the soul-hypothesis to perform the functions for which it was created without conflicting with the modern scientific view of the world. Likewise with the will, as we shall argue in this paper.

1. THE PHENOMENOLOGY OF WILLING

Nietzsche begins BGE 19 by criticizing Schopenhauer's view that 'the will alone is really known to us, absolutely and completely known', calling it an instance of what 'philosophers are in the habit of doing—he adopted a *popular* prejudice and exaggerated it'.

> Willing seems to me to be above all something *complicated*, something that is a unity only as a word—and it is precisely in this one word that the popular prejudice lurks, which has defeated the always inadequate caution of philosophers. So let us be more cautious, let us be 'unphilosophical': let us say that in all willing there is, first, a plurality of feelings, namely the feeling of the state *away from which*, the feeling of the state *towards which*, the feeling of this 'away' and 'towards' themselves, and then also an accompanying muscular sensation which, even without our putting 'arms and legs' in motion, comes into play through a kind of habit as soon as we 'will'. As feelings, and indeed many varieties of feeling, can therefore be recognized as an ingredient of the will, so, in the second place, can thinking: in every act of will there is a commanding thought—and do not imagine that this thought can be separated from 'willing', as if some will would then remain over! Third, the will is not only a complex of feeling and thinking, but above all an affect, and in fact, the affect of the command. (BGE 19)

Nietzsche's intent is to show that contrary to Schopenhauer's account of willing as something simple, 'willing is . . . something complicated, something that is a unity only as a word'. What he offers, however, is evidently a phenomenology of willing, according to which our *experience* of willing involves complexity: a variety of feelings, a ruling thought, and the affect of command. But how does the complexity of our *experience* of willing—the fact that willing is accompanied by the appearance in consciousness of feelings, thought, and affect—show that the act of willing itself is not a unity? Leiter's general claim, as we said, is that Nietzsche's phenomenology of willing functions to debunk traditional notions of the will. But this doesn't explain Nietzsche's emphasis on Schopenhauer's

commitment to the *simplicity* of the will or how he intends to counter it. We will argue that a normative interpretation of BGE 19 does a better job of explaining why Nietzsche thinks he can use the phenomenology of willing to show, against Schopenhauer, that willing itself is complex. We begin our case against Leiter's naturalistic interpretation by arguing against two aspects of his account of Nietzsche's phenomenology: his assumptions concerning (1) the kind of case for which Nietzsche offers a phenomenology in BGE 19, and (2) how much of BGE 19 is intended as phenomenology.

To understand and evaluate Nietzsche's phenomenology, we need to know what he counts as a case of willing, an 'act of will'. The widest class of cases would be actions, voluntary movements of one's body. But if Nietzsche's phenomenology is of such cases, it is quite implausible. We engage in actions all the time without much, if any, of the feeling, thought, and affect Nietzsche attributes to the person engaged in willing. Typing the word 'actions' in the above sentence was certainly an action, yet I did not—and could not, even upon reflection—detect the 'complex of feeling' Nietzsche describes here.

Nietzsche's phenomenology will be more plausible if it aims to capture a more restricted class of cases, namely, actions undertaken as the result of a deliberate decision. This is what Leiter seems to be assuming when he offers the following example to 'flesh out' the first part of Nietzsche's phenomenology.

> Sitting at the computer, I wonder whether I should go downstairs to see what the children are doing. I 'decide' to do so, and so begin to arise from my chair. I feel as if I have willed the movement. I feel the moving away from the desk and computer, the moving towards the door, and I feel the physical, or muscular movement as well. Let us call all this complex of feelings, for ease of reference, 'the bodily feelings'. (Leiter 2009: 109)

That the feelings Nietzsche refers to here are 'bodily feelings' is probably the most natural reading of the passage, although we offer an alternative below. Our main point against Leiter's computer example is that it does not help make plausible what he regards as the key element of Nietzsche's phenomenology, the 'affect of command'. It does not seem plausible that such an affect appears in consciousness whenever one decides to go downstairs to check on the children (or do anything else). Even upon reflection on such actions, it does not seem necessary to find in them anything like commanding.

The problem is even more obvious when we look at the language Nietzsche uses when he proceeds to describe this 'affect of the command' as a stance of superiority towards one who must obey:

> 'I am free, "he" must obey'—this consciousness inheres in every will; and equally so the straining of the attention, the straight look that fixes itself exclusively on one aim, the unconditioned evaluation that 'this and nothing else is necessary now,' the inward certainty that obedience will be rendered—and whatever else belongs to the position of the commander. The human being who *wills* commands something in himself that renders obedience or that he believes renders obedience. (BGE 19)

Even allowing Nietzsche some degree of exaggeration, this does not seem a plausible description of what a person experiences whenever she decides to get up from her computer and go downstairs. If this 'the drama of willing', as we'll call it, is supposed to be part of the phenomenology of willing, as Leiter must take it to be, that phenomenology is simply implausible.[3]

Nietzsche's phenomenology would be much more plausible if he has a different paradigm of willing in mind. On our proposed interpretation, Nietzsche counts as cases of willing not actions in general or actions undertaken as a result of deliberation, but actions performed in opposition to temptation. His paradigm act of will takes place in a situation of psychic conflict and struggle in which a person is faced with a choice between alternatives, one of which she is drawn to and may prefer at the moment although it flies in the face of her values, the other of which is required by her values but is not what she wants to do. We propose that overcoming temptation to act in accord with one's values—the exercise of 'willpower'—is the type of case for which Nietzsche offers a phenomenology in BGE 19. If so, the feelings that accompany 'willing' might still be grouped together as bodily feelings. But it could be equally or more plausible to interpret them as psychic feelings instead: 'the feeling of the state *away from which*' as the feeling of being in temptation, 'the feeling of the state *towards which*' as the feeling of acting or being able to act in accord with one's commitments or values. The feeling of 'away' would be the experience of the pull away from temptation, and the 'towards', the feeling of moving closer to overcoming temptation and being able to act on one's values. Either of these interpretations of the first element of Nietzsche's phenomenology fits with the 'willpower' interpretation of the passage. Further, the latter is also compatible with one of the feelings Nietzsche pays special attention to in a way that might seem to support Leiter's naturalistic reading, namely, the sensations of our muscles starting to move before they in fact do so through 'a kind of habit'. This fits at least as well as the 'willpower' interpretation: we feel our muscles moving before they do because we expect them to, that is, to do what we have just told ourselves to do.

The great advantage of taking Nietzsche to be offering a phenomenology of 'willpower' is that it can account for the necessity of both the 'commanding thought' and the 'affect of command'. As we've argued, if BGE 19 offers a phenomenology of actions (voluntary behavior) or actions undertaken as a result of deliberation, as Leiter assumes, the phenomenology is simply implausible. Even granting the necessity of a thought in the case of action based on deliberation, it need not be experienced as a command. Should we (or I) go to the movies tonight

[3] Perhaps Leiter doesn't appreciate the problem here because he translates, mistranslates, in our view, Nietzsche's 'ich bin frei, "er" muss gehorschen' as 'I am free, "it" must obey', instead of ' "he" must obey'. This does not follow any of the translations he claims to be following (and revising, if necessary), and he offers no explanation for the change. It is very difficult to see why Nietzsche would have used 'er' (instead of 'es') if he meant 'it'. There is, for instance, no masculine noun in the passage for which the masculine pronoun substitutes.

or stay home and read a book? The thought that issues in the action may well be: 'let's go to the movies', or 'I think I'll [meaning I have decided to] stay home and read'. There is no commanding thought here nor any 'affect of command'. But in a case where we exercise willpower to overcome temptation, it seems plausible that the thought that directs our behavior takes the form of a command—e.g., 'Put down the fork this minute and step away from that chocolate cake', or, slightly nicer, when one really wants to keep reading *The New York Times*: 'Time to go back to work'. And whatever the tone of the commanding thought, it must be accompanied by what Nietzsche calls the 'affect of command'. Otherwise, the person would not be telling herself what to do, but only mouthing the words, going through the motions, and she would not have the experience of her own willpower. In sum: whereas it's implausible that an affect of command belongs to the phenomenology of everyday decision-making, it is quite plausible that it must belong to our experience of exercising willpower in the face of temptation. This is the first major advantage of our interpretation over Leiter's.

A serious problem remains for Nietzsche's phenomenology, however. If it is plausible that the experience of overcoming temptation through an act of will involves an affect of command, it is (still) difficult to take seriously what we called the drama of willing that Nietzsche apparently packs into his proposed phenomenology. Recall the drama:

> What is called 'freedom of will' is essentially the affect of superiority over him who must obey: 'I am free, "he" must obey'—this consciousness adheres to every will, as does that tense attention, that straight look which fixes itself exclusively on one thing, that unconditional evaluation 'this and nothing else is necessary now', that inner certainty that one will be obeyed, and whatever else pertains to the state of him who commands. A man who *wills*—commands something in himself which obeys or which he believes obeys. (BGE 19)

Nietzsche seems to be continuing his phenomenology of willing here. If he is, 'the affect of superiority over him who must obey' is the same as the 'affect of command' that is the key component of that phenomenology. This is what still threatens the phenomenology with implausibility. When I overcome temptation through the exercise of my will, I do not seem to experience superiority over one who must obey. This is true even if the ' "he" ' or she who must obey is simply myself, as Nietzsche suggests a few lines later when he says that 'we are at the same time the one who commands *and* the one who obeys'. In exercising willpower, I am commanding myself, but this doesn't give me a feeling of superiority *over myself*. Perhaps we should take the ' "he" ' that obeys to be whatever it is that moves or threatens to move me to act contrary to my values, namely, the drives or inclinations that go against these values. The point would then be that in willing, I experience these inclinations as external to my self (since this self is constituted by my values), even though they are, of course, in another sense still *my* inclinations (they belong to me, even though I do not endorse them). But

that still does not account for the feeling of superiority over 'him who must obey' as part of the phenomenology of willing. While I may experience or treat certain inclinations as outside of myself, I do not experience them as a 'him' (or a 'her'). Nor do I feel that I am commanding my inclinations. Insofar as my feeling of command has an object, that object is myself.

Here is our solution to this problem. When Nietzsche says that what we call 'freedom of will' is 'essentially' an affect of superiority, he is no longer engaged in phenomenology. His use of 'essentially' signals that he is moving from the macro level, the level of the person and conscious experience, to the micro-level, the level of the real processes that *constitute* willing, the processes of which we are aware, whether we know it or not, when we experience willing. He is telling us how what we experience as and call 'freedom of will' is actually constituted, what it is in its essence: namely, an affect of superiority over 'him who must obey'. Who has this affect? Not (at least in the first instance) the macro-level person, but the micro-level elements, which we know Nietzsche takes to be the drives. BGE 12 claims that the soul is a 'political structure of the drives and affects', and BGE 6, that each drive (at least each one that has inspired a philosophy) 'would be only too glad to present *itself* as . . . the legitimate master of all the other drives'. BGE 19 makes explicit the image suggested by these claims, picturing some drives as commanding other drives to carry out certain actions and to ignore subversive drives who are urging otherwise. The commanding drives are presenting themselves to the other drives as having political authority, as having the authority to speak for the whole 'commonwealth', hence being the 'legitimate master' of the other drives. In taking this stance towards the other drives, they 'experience' themselves as superior to them. It is the commanding or master drive(s) who Nietzsche envisages as saying, 'I am free, "he" must obey', where the '"he"' is the drive that is being commanded to carry out the action and to ignore urgings of contrary drives. According to this account, in other words, when a person wills, commanding herself to do something against temptation, and so experiences the affect of command and 'freedom of the will', at the drive-level one set of drives is presenting itself as superior to the drives whose activity constitutes, on the person-level, the temptation of the person away from her values or commitments, i.e. from performing the commanded action.

Nietzsche presents the 'drama of willing', then, not as part of how we experience willing, but as an account of what willing is, of how willing is actually constituted. The first point in favor of this view is, as we've argued, that it prevents us from having to attribute to Nietzsche an implausible phenomenology. Further support comes from a nearby passage, which Leiter also uses to support his interpretation of BGE 19. BGE 17 argues, against Descartes, that based simply on the facts of experience or consciousnesses, we have a right to conclude only that thinking is going on, not to the existence of an 'I', a person or entity who does the thinking. 'Even the "it" contains an interpretation of the process and does not belong to the process itself' (BGE 17). Leiter quotes this passage and shows agreement

with what we have just said about it, as well as with our assumption that we are supposed to apply what it says about thinking to BGE 19's discussion of willing. But he misses or ignores what we take to be the most important implication of this point for interpreting BGE 19, namely, a denial that 'I' and 'he' (or 'it', as Leiter has it) belong to the phenomenology of willing.[4] The phenomenology involves only feelings, a commanding thought, and the affect of command. BGE 17 shows that Nietzsche is committed to saying that the 'I' and 'he' contain an interpretation of the process and do not belong to the process itself, i.e., to what appears in consciousness. Likewise in BGE 19: 'I am free, "he" must obey', is not a description of what is present in consciousness, but an interpretation of it—that is, an account of what is really going on when one experiences willing, an account of the reality that that experience reflects. That it can take account of this implication of BGE 17 is a second major advantage of our interpretation of BGE 19 over Leiter's.

2. WILLING

As we've stressed, Leiter takes the function of Nietzsche's phenomenology to be to set up a demonstration that our experience of willing misleads us about the

[4] Leiter claims that BGE 17 provides the crucial fact that Nietzsche uses to support BGE 19's denial of the will's causality, namely, that 'a thought comes when it wants, not when I want', which he takes to mean that 'our "thoughts" appear in consciousness without our having the experience of having willed them. It follows from this that one component of the experience of willing—namely, the commandeering thought—is, itself, causally determined by something other than the will' (Leiter 2009: 109). Leiter adds that, according to Nietzsche, the only will that would count as truly free—free in a sense sufficient to underwrite ascriptions of moral responsibility—is one that is *causa sui* (BGE 21). 'Since we have shown that the "commandeering thought" that is part of the experience of will is not *causa sui*, it follows that the will it helps to constitute is not *causa sui,* and thus any actions following on that experience could not support ascriptions of moral responsibility' (Leiter 2009: 115). There are a number of problems with this. First, there is no evidence in BGE 21 that Nietzsche thinks that ascriptions of moral responsibility require that the agent be *causa sui*. Second, if he did accept that requirement, it would be superfluous to argue against the will being *causa sui* on the grounds that the appearance of thoughts in consciousness is causally determined, since Nietzsche (rightly) claims that the idea of a *causa sui* is a contradiction in terms (BGE 21). Third, the fact that we can't will thoughts into existence is not a problem for any conception of free will. Consider this from the viewpoint of our claim about the paradigm case of willing that Nietzsche has in mind in BGE 19. In cases of temptation, the presence of the thought about what I should do is assumed: it defines the situation. No libertarian—much less a compatibilist—who claims that one is free to do or not do what one is tempted to do need be worried by the fact that one does not bring the thought that formulates the demands of one values into consciousness *ex nihilo*. That thought is there; one knows what it is; one also knows the competing thought: what it would be nice to do, even though it conflicts with your values. The experience of willing is presumably the experience of commanding one self to act on one of these thoughts rather than on the other. Therefore, the mere fact that one does not originally bring into consciousness by one's own will the thought that becomes one's commanding thought has no tendency to show that willing—commanding oneself to act on that thought—does not bring about actions. So if this is Nietzsche's argument for BGE 19's claim about the causality of the will, it is a puzzling and not particularly strong argument.

causality of our actions. Our claim concerning where Nietzsche's phenomenology breaks off supports a very different view of its function, such that it provides a basis for Nietzsche's theory of what willing actually is. That theory is designed to explain where the elements of the phenomenology—feelings, thought, and the affect of command—come from, to give an account of the reality reflected by the appearance of these elements in consciousness.

Leiter's account would be more plausible if he admitted this point and claimed to be offering an alternative to our account of the source of the affect of command. This alternative account can be reconstructed from the following:

> By the 'affect of the command', Nietzsche means the feeling that the thought (i.e., the propositional content, such as 'I will get up from the desk and go downstairs') brings about these other bodily feelings, i.e., of 'away from', 'towards', of, in a word, movement; and that this commanding is *who I am.* By *identifying* with the commanding thought—by taking that to be 'who I am' (on this occasion)—we feel superior, we experience this affect of superiority. *It is an identification* with the 'feeling' that the thought is superior, that it is being obeyed. (Leiter 2009: 109)

One element of this is especially strange: the idea that the commanding thought brings about the feelings of movement. Surely when we experience ourselves as willing, we believe we are bringing about the movement, not the feeling of the movement. Presumably Leiter finds himself forced to this position precisely because he thinks that Nietzsche is engaged in phenomenology, and the action itself lies outside the experience of willing. Leiter's account would be more plausible if he took Nietzsche to be breaking off his phenomenology and offering an explanation as to why the affect of command appears in consciousness, namely, that the person identifies herself with the commanding thought and believes that this thought brings about the bodily movements that constitute the action. This seems a plausible account of the passage and it might be taken to suggest epiphenomenalism if one thinks Nietzsche denies that thoughts bring about actions (say on the grounds suggested by Daniel Wegner, that the action begins before the thought occurs in consciousness).[5] But we need to consider what this 'identification' amounts to. Leiter seems to say that one takes oneself to be the commanding thought—'this commanding is who I am'. But this seems ridiculous as a piece of phenomenology. I *never* have the experience of taking myself to *be* a thought. Leiter can sensibly mean only that identification with the thought is necessary to *explain* my experiencing an affect of command. Yet this still seems highly implausible if it means that I take myself, even unconsciously, to *be* the commanding thought. There is another sense of identification, though, that offers the beginning of a plausible theory. To identify with someone is not (normally) to take oneself to be that person,

[5] Wegner, Daniel (2002). *The Illusion of Conscious Will* (Cambridge, Mass.: Harvard University Press), 34–5.

but to assume her viewpoint. So perhaps Leiter means that one experiences an affect of command when one 'wills' because one assumes the viewpoint of the commanding thought. The problem here is that thoughts themselves don't have viewpoints, although they certainly come from viewpoints, ones that belong to persons. But the viewpoint of the person, on Nietzsche's view, is the viewpoint of her drives, each of which sees the world from the viewpoint of its interests. Given this claim about the drives, it makes sense to say that the person experiences an affect of command when she 'wills' because her viewpoint in this case is the viewpoint of her commanding drives, the ones ranked higher in the political structure. It does not seem that the concept of identification adds anything that is needed here; the concept of constitution fits Nietzsche's text better. Again, the viewpoint of the person who experiences willing is constituted by, in the sense that it simply *is*, the viewpoint of the drives who use the trappings of political authority to get their way in conflicts with the other drives.

Support for this interpretation comes from a consideration of the following segment of BGE 19:

> But now observe the strangest thing of all about the will—about this so complex thing for which people have only *one* word: inasmuch as in the given circumstances we are at the same time the commanding and obeying one, and as the obeying one know the sensations of constraint, compulsion, pressure, resistance, motion which usually begin immediately after the act of will; inasmuch as, on the other hand, we are in the habit of disregarding and deceiving ourselves over this duality by means of the synthetic concept 'I'; so a whole chain of erroneous conclusions and consequently of false evaluations of the will itself has become attached to willing—to such a degree that one who wills believes sincerely that willing suffices for action. (BGE 19)

The 'strangest thing about the will' consists in the conjunction of two facts: first, that the will admits of a duality and, second, that we disregard and deceive ourselves about this duality 'by means of the synthetic concept "I" '. Leiter glosses this as follows:

> In other words, talk of 'I'—as in '*I* will go downstairs'—obscures an elemental fact: it's my body that is 'obeying' my will, so 'I' am also one who obeys, as well as one who commands. Of course, we don't experience it, or think of it, that way: we identify the 'I' with the feeling of *commanding*, not the feeling of obeying. (Leiter 2009: 110)

But this isn't quite right. Nietzsche doesn't claim that my use of 'I' obscures from me the fact that I also obey—he takes for granted that we recognize this fact. His claim is that what we call 'the will' is a 'complex thing for which the people have only one word'. The will is 'complex' in the sense of having different components, a part that commands and a part that obeys. This is what is obscured by our use of 'I' to talk about both the commanding party and the obeying one: that the will has different components, not that I obey as well as command. And it is not even that we can't think of ourselves as

also obeying while we are willing. To the extent that I think of myself as commanding at all, I think of myself as commanding myself—and certainly not some 'him' or 'her', as we have already argued—and therefore can easily recognize, even if I don't usually have it in the forefront of my consciousness while willing, that I am also obeying myself. Yet, insofar as we experience ourselves as willing, Nietzsche is claiming, we assume the viewpoint, and thus experience the affect, of 'something in us' that commands, and not the viewpoint of 'something in us that obeys'. Now what is it in us that commands and obeys? The obvious answer is the drives. The dominant drives in the political order of the drives—those that have the authority to speak for the whole commonwealth—command. These are the drives whose high political ranking constitutes a person's most basic commitments. The obeying drives are those that carry out the command, the dispositions to basic actions, we might say, as well perhaps as the subversive drives who aim to tempt the person away from acting on her values.

As further evidence for our interpretation, consider Nietzsche's reformulation of 'freedom of will' later in the passage:

> 'Freedom of will'—that is the expression of the complex state of pleasure of willing that commands and at the same time identifies itself with the one who carries out the command; that, as such, enjoys also the triumph over obstacles, but thinks within himself that it was really his will itself that overcame them. In this way, the one who wills adds the feelings of pleasure of its successful executive instruments, the useful 'underwills' or under-souls—indeed, our body is but a social structure composed of many souls—to the feeling of pleasure as commander. (BGE 19)

Notice, first, that Nietzsche's equation of 'underwills' and 'undersouls' implies that the will is the soul. Because, as we have noted, Nietzsche takes the soul to be the 'political structure of the drives and affects' (BGE 12), a structure of legitimacy and authority, not simply strength, this confirms our claim that the parts of the will are the drives and that the structure that allows them to form the will is political. But if the will is the soul, how can it be composed of 'many souls'? The point seems to be that the whole (will or soul) is composed of parts (the drives), each of which is a political structure of many other drives. For instance, Nietzsche presents the drive to knowledge as the product of a long historical and cultural process, which brought many different drives together under a common project (GS 110; cf. GM I: 1). Second, the drives that carry out the orders of the politically dominant drives appear to be dispositions of the body—hence, as we've said, dispositions to carry out basic actions, such as moving one's feet. The third and most important point is that what constitutes our experience of 'freedom of will' combines elements from both components of the will: the viewpoint and therefore affect of superiority of the *commanding* drives, but also the pleasure of the *commanded* drives insofar as they overcome obstacles in executing the order.

The section of the passage from which we draw this last point continues as follows:

> *L'effect c'est moi*; what happens here happens in every well-constructed and happy commonwealth; namely, the governing class identifies itself with the successes of the commonwealth. (BGE 19)

The 'governing class' in the case of the will can only be the commanding drives: those ranked highest in a political order. They do not actually execute the bodily movement they have commanded; such movement is carried out by drives or habits related specifically to such movement. But the 'governing class' formed by the commanding drives 'identifies itself with the successes of the commonwealth' in the sense that it takes credit for what the commanded drives have done, which is what happens in 'every well-constructed and happy commonwealth'. The pharaoh says that he built the pyramids, when, more literally formulated, he merely ordered others to do so.

So it is the commanding drive that 'identifies itself with the one who carries out the command'. But because the person's experience of willing is the appearance in consciousness of the viewpoint of the commanding drives, and these drives identify with the 'successes of the commonwealth', the person also experiences the pleasure of 'the one who carries out the command'. 'In this way', Nietzsche tells us, the person adds 'to the feeling of pleasure as commander' 'the feelings of pleasure of its [the commander's] successful executive instruments'. So, for example, when I get myself to walk up the hill by telling myself that I could use the exercise even though I don't want to, I get the pleasure not only of commanding the action but also of 'triumph[ing] over obstacles', which I get by obeying, i.e., by walking up the hill. The upshot is that to have freedom of the will is to be a well-constructed and happy commonwealth of drives, one in which the drives with political authority command and are obeyed by the drives that actually do the work. The politically superior drives rightly take credit for the work done by the other drives, because it would not have been done in the circumstances, if they had not commanded it.

This way of putting it may be misleading, however, for we do not hold that willing is required for all action. As we interpret Nietzsche, the will is causally connected to action even when there is no commanding going on, as indeed there is not in the majority of human action. The will is the political order of the drives, an order of legitimacy, but in 'a well-ordered and happy commonwealth' (BGE 19), the causal order of the drives, an order of strength, is largely in accord with the political order, and this is no accident. At the very least, the causal order wouldn't be as it is if it were not in accord with the political order—i.e., with the person's values—and, in at least some cases, it is as it is precisely because of the political order. In these latter cases, it is precisely because the authority of the commanding drives have been recognized that certain practices have been undertaken (sometimes in response to commands, sometimes simply

because the authority is recognized) so that, as a result, the drives with political authority are now also causally stronger than those that might otherwise try to subvert them, and need not resort to commands to have the will influence action. This explains Nietzsche's otherwise puzzling claim in *Ecce Homo* that he knows nothing of willing from experience. ' "Willing" something, "striving" for something, envisaging a purpose, a wish—I know nothing of this from experience' (EH I: 9). Given Leiter's interpretation of BGE 19, it is difficult to understand Nietzsche's denial that he has had any experience of willing. We can grant that Nietzsche may be exaggerating here, but his denial would be much more than exaggeration if he thinks that the experience of willing discussed in BGE 19 is something that accompanies all action taken in response to a deliberate decision. On the other hand, if our analysis is correct, and the experience of willing detailed in BGE 19 is the experience of willpower, it makes perfect sense for Nietzsche to deny that he knows this from experience (exaggerated though this may be) in a passage devoted to advising us that 'the whole surface of consciousness . . . must be kept clear of all great imperatives. Beware even of every great word, every great pose' (EH I: 9).

We conclude this section with a discussion of the ending of BGE 19, which, as mentioned, seems particularly difficult to interpret. Although our several initial reactions to this segment was 'what in the world is he talking about?' we now think that the interpretation of BGE 19 we have developed provides a framework for understanding it. We do not see how it can make much sense given Leiter's naturalistic interpretation of the passage, and suggest that Nietzsche placed it at the end of BGE 19 precisely to give readers pause about the naturalistic reading to which he knew the rest of the passage would incline and perhaps seduce them. Here is the ending:

> In all willing, it is absolutely a question of commanding and obeying, on the basis, as already said, of a political structure composed of many 'souls': on which account a philosopher should claim the right to include willing in itself within the field of morality: that is, of morality understood as the theory of the relations of dominance under which the phenomenon 'life' arises. (BGE 19)

Note that Leiter cannot take Nietzsche to be talking here about the phenomenology of willing because he denies that obeying is part of that phenomenology. Nietzsche can only be talking about willing itself here. Having told us in BGE 12 that the soul is a political structure of the drives and affects, Nietzsche is now talking about how that political structure works to bring about actions: through the commanding and obeying of the drives. We discuss objections to the plausibility of that theory below, but even if it is implausible, it seems to be the one that Nietzsche puts forward in the passage.

But what can it mean to say that philosophers have a right to include willing 'within the field of morality', much less to call morality 'the theory of the relations of dominance under which the phenomenon "life" arises'? The

naturalistic interpretation of willing contains no resources for explaining these claims—nothing that connects willing to morality, or that seems able to explain how life comes about through 'relations of dominance'. We can begin to make sense of these claims if we recognize that Nietzsche takes willing to be essentially connected to values.

Nietzsche's characterization of morality in terms of 'relations of dominance' that give rise to 'life' requires us to ask what he means here by 'life' and by 'morality'. As to the first, we have argued elsewhere that BGE 9 defines 'life' in normative rather than biological terms, as equivalent to 'true life', to living in accord with values or ethical commitments, and that this is how he uses 'life' throughout the first part of BGE (Clark and Dudrick, forthcoming). We suggest that Nietzsche is doing the same thing here: 'life' is equivalent to 'true life', a life lived in accord with ethical values. Further, such values are precisely what Nietzsche has in mind when he speaks of 'morality' here; he is using the term in the wide sense that makes it equivalent to 'ethics'—the narrow sense being precisely the kind of morality that masquerades as morality itself and that Nietzsche's philosophy 'beyond good and evil' aims to overcome (BGE 32; Clark 2002). With these two claims in place, we can see that the 'relations of dominance' are the political relations among the drives that constitute a person's having ethical values, and are therefore that in virtue of which she has 'life' in the normative sense. Being thus able to offer a plausible interpretation of the end of the passage adds to the advantages of our interpretation of BGE 19 over Leiter's.

3. THE ERROR ATTACHED TO WILLING

Although we have offered strong evidence that BGE 19's phenomenology of willing seeks to *rehabilitate* rather than to debunk the traditional notion of willing, we recognize that BGE 19 rejects *something* attached to that notion. Our aim in this section is to explain what this is in a way that leaves the causality of the will intact.

BGE 19 claims, in fact, that 'a whole chain of erroneous conclusions and consequently false evaluations of the will itself [des Willens selbst] has become attached to willing [Wollen]', explaining that because 'the synthetic concept "I" ' leads us to disregard the duality of the will, we are misled into believing '*that willing suffices for action*', i.e., that acts of will are causally sufficient to bring about actions. Leiter takes Nietzsche's point here to be that 'the phenomenology of willing, no matter how vivid, does *not* in fact mirror or reflect or . . . track an actual causal relationship'.[6] This is because 'the commandeering thought,

[6] Leiter adds a parenthetical qualification here, suggesting that BGE 19 may be denying only that the phenomenology of willing tracks 'a causal relationship *sufficient to underwrite ascriptions of*

with which we identify because it gives us a feeling of superiority, is not in fact identical with anything that actually stands in a causal relationship with the resultant action' (Leiter 2009: 111). We make the error of thinking that it is, according to Leiter, because we mistake a relationship of *correlation* for one of *causation*. Leiter takes this to be the point of Nietzsche's claim that because

> in the great majority of cases willing takes place only where the effect of the command, therefore obedience, therefore the action, was to be *expected*, the *appearance* has translated itself into the feeling, as if there were here a *necessity of effect*. (BGE 19)

But note that this passage makes explicit that 'obedience, therefore the action', is the 'effect of the command', and that this commits Nietzsche to the claim that willing (commanding) does cause action. Leiter can say that Nietzsche must mean only that the action is the 'apparent effect of the command,' but there is no reason to interpret it that way other than the needs of Leiter's interpretation. If his point is that the experience of willing is merely correlated with the occurrence of the action but is not its cause, it is strange that Nietzsche would include in the sentence, 'the effect of the command, therefore obedience, therefore the action'. The natural way to read the passage, given how Nietzsche actually wrote it, is as taking for granted that willing (commanding) causes action, and seeking to explain how we come to assume 'that willing *suffices* for action'. Calling this assumption 'erroneous' is not a denial that willing bears a causal relationship to action. Nietzsche denies only that willing is *sufficient* to produce action.

But does this make sense? Can we affirm that A causes P while denying that A suffices for P? To defend an affirmative answer to both questions, we invoke J. L. Mackie's gloss on causation: 'The statement "*A* caused *P*" often claims that *A* was necessary and sufficient for *P in the circumstances*' (Mackie 1993 [1965]: 39, emphasis added). To say, for example, that the faulty brakes caused the accident is to say that the faulty brakes were necessary and sufficient for the accident *given the circumstances*—given that the car was moving at a certain speed, for example.

moral responsibility' (Leiter 2009: 108; emphasis added). But this qualified interpretation is really a very different claim than the one BGE 19 seems to be making about the causality of the will. Leiter was right to see in that passage the suggestion that willing does not cause action at all, that it is epiphenomenal. It is this suggestion—made by Nietzsche's text and defended in Leiter's account of it—that we are concerned to rebut in this paper. The qualified interpretation—that because willing is determined, we are not morally responsible for our willed behavior—is a very different matter, and requires a different response. Although a full response is beyond the scope of this paper, we note that it is difficult to see why Nietzsche would think the phenomenology of BGE 19 is needed for that claim (cf. note 4 above). Further, what Leiter would need to make attributing this qualified claim to Nietzsche plausible is to show that BGE 21 assumes, as he claims, that all forms of compatibilism are false. But he offers no evidence for this, and, on the face of it, BGE 21 is an argument for compatibilism. Far from rejecting ascriptions of responsibility as requiring the nonsensical notion of a *causa sui*, BGE 21 actually criticizes such a rejection as stemming from an erroneous understanding of causality. We take up these matters in our forthcoming book.

One might consistently deny that (a) the faulty brakes *sufficed* for the accident without thereby denying that (b) the faulty brakes *caused* the accident, if (a) is interpreted as neglecting the fact that *other* conditions were necessary (e.g., the slippery condition of the road and the high rate of speed at which the car was moving). Nietzsche does something similar in the case of willing: he denies (a) that 'willing *suffices* for action' without denying (b) that willing causes action, because he takes (a) to neglect the fact that *other* conditions are necessary to allow willing to bring about the action. What are these conditions? Well, consider that a person's willing *is constituted by*—and so is identical with—the high-ranking drives commanding lower-ranked drives to do something—whereas the person's acting is constituted by the commanded drives obeying these commands. The further conditions to be met if willing is to bring about action are therefore whatever conditions must be in play for the commands of the superior drives to bring about the obedience of the commanded drives. The main condition would presumably be that the commanded drives recognize the authority of the commanding drives, that the commanded and commanding drives thus exist in a 'well-constructed and happy commonwealth' (BGE 19). What makes this possible is an important theme of Nietzsche's work, but not one that he addresses in this passage.

The error of thinking that 'willing *suffices* for action', then, is not that of thinking that willing causes action, but of thinking that commanding an action is enough, all by itself, to bring the action about. We are misled into thinking that it is enough because the experience of commanding an action is usually followed by the action; this is because usually the commanded drives obey the commands of the politically high-ranking drives. This is actually a conceptual necessity. If the commands of the politically high-ranking drives *weren't* usually obeyed, these drives would not be politically high-ranking, and nothing they could do would count as 'commanding', just as the behavior of the commanded drives would not count as recognizing their authority or obeying. We therefore have a tendency, one that bolsters our sense of power (because we are assuming the viewpoint of the commanding drives), to think that willing (commanding) brings about action all by itself, thus overlooking everything that conditions the commanded drives' willingness to recognize the authority of the commanding drives, and thus to obey these commands.

But is this error so important that Nietzsche would put it forward as his major problem with traditional accounts of the will? We think it is for a number of reasons. First, it renders it difficult to recognize what is necessary to actually bring about changes in one's self or life. Consider what happens if one thinks that willing does suffice for action. One thinks that one can affect the causal order of one's drives—hence what one does—simply by issuing commands, and forgets that commanding can do nothing by itself; that commanding is effective only if one's drives exist in a 'well-constructed and happy commonwealth', hence that bringing about such a commonwealth of drives is what is necessary if one

is to change one's life. Second, this provides an excuse for the mistaken belief in 'freedom of the will' in what Nietzsche calls the 'superlative metaphysical sense' (BGE 21). Ignoring that fact that willpower only brings about action if the commanded drives are willing to obey, allows one to ignore all of the moral luck—the influence of 'the world, ancestors, chance, and society' (BGE 21)—that goes into having one's drives exist as a 'well-constructed and happy commonwealth' (BGE 19)—and thus to believe that one has total causal responsibility for one's actions. Third, it is precisely susceptibility to this error that makes the naturalist likely to reject the very existence of the will. Consider that Leiter appears to reason as follows:

1. If there were a will, then willing would suffice for action.
2. Willing does not suffice for action.
3. Therefore, there is no will.

On our view, Nietzsche accepts 2 but denies 1. He holds that the will is the normative order of the drives (i.e., the soul) insofar as that order is involved in action, and that willing is the commanding of the normatively high-ranked drives. Why would it be problematic for Nietzsche if he had to reject the will itself and not just a particular theory of what willing involves? As we see it, this would render him unable to accomplish a key aim of his psychology: to account for the fact that human beings are capable of acting morally or categorically—that they are, that is, capable of doing something not because they desire to or find it in their interests, but because they take it to be good.

When a person acts on the basis of categorical values, the causally effective drive is the one that is also highest in the political ranking—and such an action is therefore reflective of her will. As we've argued, BGE 19's analysis of willing concerns a case in which a person is inclined to act in a way that contradicts her values, which means that the drive that is causally strongest at the moment is not the one that is highest in the political order of the drives. When willing (commanding) is successful, the drives' causal order is brought into line with their normative order through the use of political authority, so that the person's will (the normative order of her drives) is reflected in her action.

4. TWO KINDS OF PSYCHOLOGICAL THEORY

We've argued that the account of willing Nietzsche offers in BGE 19 equates the will with the soul, and that his theory of the soul provides a basis for distinguishing what a person values from what she merely desires. But having a will is not the same as willing. On our reading of Nietzsche's view, a person *wills* when she tries to overcome temptation and to act in accord with her values. Such willing is, according to BGE 19, constituted by a drive's commanding some other drives. It is just in case the drives commanded *obey* that the person's

willing is successful. To will is thus, on Nietzsche's view, to exercise what it commonly called 'willpower'—an exercise which, on the level of the drives, is constituted by a drive's attempt to modify the causal order by way of its standing in the political order. On our view, mentioned at the outset and developed in our forthcoming book, Nietzsche takes this political order—the person's soul or will—to constitute the person's values. And because a person usually acts in accord with her values—if she *didn't*, they wouldn't *be* her values—the political order and the causal order (which issues in action) are, for the most part, the same. As a consequence, a person's actions will usually *express her will* without her having to *engage in willing*—without her having to exercise willpower. Such a person is thus constituted by a 'commonwealth' that is, for the most part, 'well-constructed and happy' (BGE 19).

While the text supports our attribution of the position described here to Nietzsche, the principle of charity may seem to militate against it. That is, while Nietzsche surely *seems* to be saying that the drives form a political order, surely he can't *mean* that literally—to do so would seem to be to treat the drives as proto-persons. We are, one might contend, better off losing the will or soul than keeping one composed of homunculi. Two questions, then: first, doesn't our reading have Nietzsche regarding the drives as homunculi? And second, doesn't this render the view it attributes to Nietzsche ridiculous? In what follows, we'll argue that while the answer to the first question is a qualified Yes, the answer to the second is No.

In order to answer the charges presented here, it is helpful to consider Daniel Dennett's characterization of the 'two broad strategies one could adopt' in pursuing psychology, the 'bottom-up' and 'top-down' strategies. He tells us that 'a *bottom-up* strategy'

> starts with some basic and well-defined unit or theoretical atom for psychology, and builds these atoms into molecules and larger aggregates that can account for the complex phenomena we all observe. (Dennett 1978: 110)

This strategy takes some non-intentional unit (stimulus-response pairs or neuron signals, for example) and attempts to show how such units can be understood as to constitute intentional, psychological phenomena like human beings acting for reasons. The second strategy is the '*top-down* strategy', which

> begins with a more abstract decomposition of the highest levels of psychological organization, and hopes to analyze these into more and more detailed smaller systems or processes. (Dennett 1978: 110)

This strategy takes an intentional phenomenon like that of a human being willing an action and attempts to show how it can be understood in terms of simpler—though still intentional—phenomena. On the reading we've offered, Nietzsche's account of the will is clearly an instance of the 'top-down' approach. He attempts to explain a person's willing in terms of a will constituted by drives

which themselves exhibit agency of a sort, i.e., that involved in commanding and obeying.[7]

We can develop our account of Nietzsche's position here by responding to an apparent dilemma. One might claim that what we've presented as Nietzsche's explanation of acts of will (i.e., cases in which a person resists temptation and acts on her values) is either not an *explanation* or it is not *of* such acts of will. To see this, consider that either we have Nietzsche treating the drives as persons or we do not. If we *do* have him treating the drives as persons, then his position is problematic not just because it's ridiculous to regard drives in this way, but because we have him explaining the activities characteristic of persons in terms of drives—which engage in the activities characteristic of persons. In that case, his is not an *explanation* in any recognizable sense. However, if we *don't* have him treating the drives as persons, then the talk of 'commanding' and 'obeying' in BGE 19 seems to fall away, and we're left with a view according to which it is simply the strongest drive that produces behavior. In that case, his account has explanatory power, but rather than illuminating 'acts of will', it explains them away: what *appear* to be cases of resisting temptation and acting on values turn out to be indistinguishable, on the level of the drives, from cases of a simple indecision followed by an action. Therefore, Nietzsche's explanation of acts of will is either an *explanation* or it is not of such acts of will, but it is not both.

The first horn of the dilemma says that if our reading has Nietzsche treating the drives as homunculi, then it cannot take him to have offered an explanation of the agency characteristic of human beings. To see why this isn't so, it's helpful

[7] While Dennett tells us that the process of 'top-down' explanation continues 'until finally one arrives at elements familiar to the biologists' (Dennett 1978: 110), it's not clear that Nietzsche would agree. The reason is that the last part of the strategy is ambiguous in an important respect. It might be that to 'arrive' at 'elements familiar to the biologists' is to account for the *existence* of the intentional phenomenon in question using these 'elements'. So understood, the 'top-down' strategy aims to explain how it is that beings who, e.g., think, desire, and will came into existence via processes that can be understood from a naturalistic perspective. But it might also be that to 'arrive' at these elements is to use them to account for the *intentionality* of the phenomenon in question. Understood this way, the strategy attempts to reduce meaning to causal processes. Put another way, where the first version of the strategy attempts to understand how the 'space of reasons' arose from the 'space of causes', the second version attempts to understand what goes on in the 'space of reasons' in terms of the space of causes.

This distinction is important, since on the reading we advance in 'The Naturalisms of *Beyond Good and Evil*' and in our forthcoming book, Nietzsche cannot accept the second version of this strategy. More specifically, our reading of GS 373–4 contended that in those sections Nietzsche puts forward the claim that a complete causal (or 'mechanistic') account of the world is not a complete account of the world—such an account will not, that is, provide access to all truths. This is not to deny that all phenomena can be understood in causal terms; it is only to claim that there are true descriptions of the world that cannot be put in purely causal terms. Among these true descriptions are those involving inquiry—thus, one who, like Mr Mechanic, claims that all phenomena can be understood in causal or mechanistic terms finds himself unable to make sense of his own status as an inquirer. This is what makes his 'among the stupidest interpretations, one most lacking in meaning' (GS 373).

first to note that a psychological analysis is not viciously circular simply by virtue of being 'top-down'. As Dennett puts the point,

> Homunculi are *bogeymen* only if they duplicate *entire* the talents they are rung in to explain. If one can get a team or committee of *relatively* ignorant, narrow-minded, blind homunculi to produce the intelligent behavior of the whole, this is progress. (Dennett 1978: 123)

The question, then, is whether the activities in which the drives engage are *simpler* than the activities of the person which they are 'rung in to explain'. And it's clear that they are. For recall that we take BGE 19 to be Nietzsche's account of what's going on when a person engages in an act of will, i.e., when she resists temptation and acts on her values. The account does not have *the drives* resisting temptation and acting on values—it has them simply commanding and obeying. Because these activities are simpler than the one for which they are supposed to account, there is no problematic circularity here.

But if making the activities of the drives relatively simple helps to overcome the first horn of the dilemma, it might seem to run into the second. For one might wonder whether Nietzsche's account doesn't really amount to saying that the strongest drive wins out in the end. If that's so, then it doesn't offer the basis for distinguishing between cases of temptation (which involve values) and cases of simple indecision (which need not).

To see why Nietzsche's account *can* distinguish between temptation and indecision, note first that there is a trivial sense in which the strongest drive always wins out, namely, if strength is defined in terms of its ability to win out in case of conflicts. But in any other sense Nietzsche's account has as its intention to *deny* that human behavior is simply a matter of the strongest desire winning out: that is precisely why he has the drives commanding and obeying in certain cases. When one acts in accord with one's values rather than one's momentary desires, the commands of some drive(s) counteracts and overcomes the mere physiological strength of the drives. In such a case, the drives aren't simply exerting brute causal strength (as they do when a person is in a state of indecision); rather, on Nietzsche's account, one of the drives is exerting *political* authority.

The dilemma proposed above, then, is a false one: the view we've attributed to Nietzsche is indeed an explanation of what goes on when a person engages in an act of will. One might admit this, however, and still contend that while it is *potentially* an explanation, it is actually quite ridiculous. For the view does, after all, have *the drives* forming a *political order*. But how in the world are the drives supposed to have done this? How—as we might imagine Nietzsche himself to ask—could such a thing come to be, given that the development of human beings is continuous with the rest of the animal world?

Nietzsche recognizes the need to answer this question and he does so, not in BGE but in the work Nietzsche calls 'a supplement [to] and clarification [of]' that work, *On the Genealogy of Morality*. On our view, Nietzsche takes the political order of the drives to be that in virtue of which human beings have values and so are persons; this being so, we should be able to see how this political order comes to be by examining Nietzsche's account as to how *homo sapiens* became persons. And this is precisely what we find in GM: like Plato, Nietzsche uses reflection on the state as a way of understanding the soul. Unlike Plato, Nietzsche gives an account of the *origin* of the soul, one that shows his understanding of it to be consistent with naturalism. He does this by showing first that the origin of the state is clearly explicable in naturalistic terms. He then shows how the origin of such a political order *among* human beings illuminates the origins of such an order *within* human beings—thus showing that one need not invoke anything that is not part of the natural world in explaining the advent of the soul.

We must leave for our forthcoming book an account of *how* Nietzsche uses the origin of the state to illuminate the origin of the soul. We end this paper by giving reason to think *that* he at least intends to do so. We offer as evidence GM II: 18. There Nietzsche claims that

> The active force that is at work on a grander scale in those violence-artists and organizers that builds states, is basically the same force that here—inwardly, on a smaller, pettier scale, in a backwards direction, in the 'labyrinth of the breast', to use Goethe's words—creates for itself the bad conscience and builds negative ideals: namely that *instinct for freedom* (speaking in my language: the will to power).

Here Nietzsche claims that the same will to power expressed by the blond beasts also works 'inwardly, on a smaller, pettier scale' or level—it works, that is, not just on the level of the person but on the level of the drives. He tells us that as the blond beasts expression of the will to power produced 'the state', what does its expression among the drives produce? The answer, Nietzsche tells us, is 'the bad conscience' and 'negative ideals'—both of which, we argue, are instrumental in making human beings persons precisely because they are so in producing a political order among the drives.[8]

[8] One might claim that, despite what Nietzsche seems to intend, drives are not the *kind of thing* among which political order could obtain: drives are behavioral dispositions, not potential citizens. Here we should note that an order and its activity need not be sophisticated or at all high-minded in order to be political. For while acting in accord with values in the face of temptation is found only among human beings, the kind of political life we take Nietzsche to attribute to the drives here is common throughout the animal kingdom. 'Dominance hierarchies'—in which the rank order and not merely the brute strength of members determines their access to various resources—are found among chimpanzees, wolves, and other social animals.

REFERENCES

WRITINGS BY NIETZSCHE

[1886]. *Beyond Good and Evil*. trans. R. J. Hollingdale (New York: Penguin Classics, 2003).

[1886]. *Beyond Good and Evil* in *Basic Writings of Nietzsche*. trans. W. Kaufmann (New York: The Modern Library, 1966).

[1886]. *Beyond Good and Evil*. trans. J. Norman (Cambridge: Cambridge University Press, 2002).

[1882–1887]. *The Gay Science.* trans. Josefine Nauckoff (Cambridge: Cambridge University Press, 2001).

[1887]. *On the Genealogy of Morality.* trans. Maudemarie Clark and Alan Swenson (Indianapolis, Indiana: Hackett Publishing Company, 1998).

OTHER WORKS CITED

Blackburn, Simon (1993). 'Hume and Thick Connexions,' *Essays in Quasi-Realism* (Oxford: Oxford University Press), 94–107.

Clark, Maudemarie (2002). 'On the Rejection of Morality: Williams' Debt to Nietzsche', in R. Schacht (ed.), *Nietzsche's Postmoralism* (Cambridge: Cambridge University Press), 100–22.

____ and Dudrick, David (2005). 'The Naturalisms of *Beyond Good and Evil*', in K. Ansell Pearson (ed.), *A Companion to Nietzsche* (Oxford: Blackwell Publishing), 148–67.

____ (Forthcoming). *An Introduction to Nietzsche's* Beyond Good and Evil: *The Magnificent Tension of the Spirit*.

Dennett, Daniel (1978). 'Artificial Intelligence as Philosophy and as Psychology', *Brainstorms* (New York: Bradford Books), 109–26.

Leiter, Brian (2009). 'Nietzsche's Theory of the Will', (this volume); first published in *Philosophers' Imprint*, 2007, vol 7: 7, 1–15.

____ (2002). *Routledge Philosophy Guidebook to Nietzsche on Morality* (London and New York: Routledge).

Mackie, John L. (1993 [1965]). 'Causes and Conditions', in E. Sosa and M. Tooley (eds.), *Causation* (Oxford: Oxford University Press), 33–55.

Index

Acampora, Christa Davis 204–5, 209–10
acting 62,108–111, 134, 135, 139, 156, 187–9, 193, 194 (n 11);
  contrasted with doing p. xv, 33–4, 39–44, 119 (n 16), 205
Aesthetic 9
  -evaluation 162, 168
  -experience 82, 172–3, 214–7
Agency 4, 33–7, 39, 40, 64, 80, 91, 118, 121, 122, 127 (and n 1), 132, 135, 137 (and n 22), 138–49 (and n 34 on 143), 155–6, 159, 169, 184, 191, 202, 208–12, 215–17, 265
  -free will 34–7, 39, 41, 44–7, 119 (n 16), 205
  -free will, contrasted with deserts free will 38, 39 (and n 8), 40, 47, 60 (n 21), 205
  -as an achievement 41–7, 57, 61, 66, 79
Aristotle 76, 194 (n 111), 190 198–202, 213, 219
affects; *see also* drives; instincts; interests. 26, 51–64, 66, 70, 90, 104, 109, 111, 113–116, 118 (n 13), 125, 133, 134, 137, 141, 143, 155–71, 177, 249, 252–7
affirmation xvii, 8, 10, 81–5, 83 (n 44), 97–8, 169, 224–6, 236, 242–3
amor fati 97, 98, 153 (n 4), 213–20, 227.
ascetic ideal xvii, 23–7, 90, 97, 102–5, 139, 140, 144, 170–2, 224, 228–30.
autonomy 10, 11, 33–4, 37–8, 41, 47, 60–1, 65, 90, 94, 111 (n 16), 128, 142, 156–7, 162, 164, 192, 194, 202–6, 210–12, 217, 219, 220

bad conscience 54, 66, 142, 233, 238–9, 267
'becoming what one is' 92, 210
Blackburn, Simon 159–60

causa sui 40, 41, 62, 63, 69, 113, 114, 121, 124 (n 21), 205, 254 (n 4)
character 38, 47, 53, 58, 62, 66, 77, 78, 79, 93, 154, 165, 169, 211, 214, 215 (n 28), 216, 227, 242–43.
  -as stable unity of drives 125, 140, 169
  -friendship 198–200
  -'giving style to one's' 77, 93, 214–16, 227

compatibilism 40, 76 (n 21), 107, 115
  see also free will
Cox, Christoph 54 (n 6), 55 (n 10).
consciousness 6 (n 8), 10, 48, 92, 112, 118 (n 14), 119, 133, 136, 138–42, 155 (n 7), 177, 204, 233–4, 247, 253–9;
  as epiphenomenal 3–4, 118 (n 14), 247;
  self consciousness 3–4, 10, 13, 58, 89, 157, 161

Deigh, John 237
Dennett, Daniel 4, 129, 142, 264, 265 (n 7), 266
Descartes, Rene 46, 175, 253
determinism 33, 36–8 (n 5), 63, 70, 73, 107, 114 (n 11), 205
drives 55–9, 90–5, 105, 124, 132–49, 152, 154, 161, 256–9, 262–7;
  the Self and 4, 13 (n 18), 248;
  structure of 37–8, 42–8, 132;
  master 42, 253;
  see also affects; instincts; interests

ego ('I') 2–14, 48, 57–9, 118, 138, 141–3, 231–2, 254–8;
  drives in relation to 4, 13 (n 18), 55, 57, 59, 90–5
egoism 9–10, 233;
  moral 65–66, 74;
  psychological 197–200, 219
Epiphenomenalism 108, 113 (and n 9), 115, 117–23;
  of the will 247–8, 254–9, 260 (n 6);
  of consciousness 3–4, 48 (n 19), 136
emotions 52, 158, 160–8, 200, 228, 47 (n 17)
eternal recurrence 9, 14, 20, 213, 223–43;
  as a test of freedom 97–9;
  Freud and 224 (n 5), 228;
  see also amor fati
expressivism 158

fate 62, 83, 86, 89, 103–5, 113–114 (n 11), 148, 186–7, 189–190, 204–8, 213–220;
  willed in eternal recurrence 96–9, 227;
  see also amor fati
Fichte, Johann 10–13, 39 (n 8)
flourishing 20, 90, 94–6

freedom 9–12, 35–7, 43–8, 60–6, 69–71, 73–82, 84–6, 89–93, 105, 127–49, 152–7, 161–2, 169–171, 177, 187–94, 203–5, 209, 210, 212, 252–8, 263;
as a capacity 132, 135;
exemplified in the sovereign individual 36–7, 128–9, 142–3;
as an achievement 60, 152;
and nihilism 96–9
free will 4, 33–48, 60–6, 69–72, 91, 107–25, 127, 141–3, 151–2, 204–5, 254 (n 4);
agency contrasted with deserts 35, 39–41, 60 (n 21), 205;
see also compatibilism; fate; freedom; will
free spirit 43–4, 74, 85, 89, 94 (n 17), 104, 147–8, 151–6, 170–7
Freud, Sigmund 13, 27–30, 48, 109, 115, 118 (n 14), 223–41

Gardner, Sebastian 48 (n 20), 59, 92 (n 11), 230
genealogy 70, 75, 131–2, 145–9
Gemes, Ken 60 (n 21), 76 (n 21) 119 (n 16), 205
German Materialists 115
German Romantics 13, 20 (n 30), 78
God, death of xvii, xiv, 23, 28, 41, 61–5, 96–102, 144, 226–30, 239
Goethe, Johann Wolfgang von xiv, 34, 35, 42–5, 62, 93, 95, 98, 103, 145, 148
Green, Michael Stephen 22 (n 33)
Guay, Robert 156 (n 8), 211
guilt 23, 35, 38–41, 54, 65–6, 129, 140–4, 223–5, 233, 236–41
greatness 94, 152–3;
see also higher men; freedom

Habermas, Jürgen 19 (n 28), 21–2
happiness 197–201
Hatab, Lawrence 205
health 47, 49, 66, 75, 135, 145
Hegel, G.W.F. 39 (n 8), 80, 81, 219
Heidegger, Martin 83 (n 44), 176
higher men 94, 224, 241–3;
see also greatness; types
Hume, David 5, 12, 15–17, 37–40, 46 (n 16), 160–1
Husserl, Edmund 167 (n 25)

instincts *see also* affects; drives; interests 52, 55 (n 10), 65–8, 129–30, 147–8, 231, 233, 237–40
interests *see also* affects; drives; instincts 4–6, 21, 71, 80, 135–49, 176–7, 256

Janaway, Christopher 72 (n 9), 127 (n 2)

Kant, Immanuel 1, 5, 7–13, 19–21, 35, 39, 46–7, 52, 57–8, 69, 119 (n 18), 113 (n 11), 131, 155 (n 7), 159, 164, 171 (n 29), 192–4 (n 11), 197–202, 210, 219–20, 230
Katsafanas, Paul 113 (n 9), 118 (n 14), 129 (n 6), 133 (n 15), 139 (n 27)
knowledge 18, 22 (n 33), 75 (n 18), 82, 176;
and 'eternal recurrence' 229–30;
perspectivism and 51–4, 59–60, 155, 230;
role of affects in 51, 55

Lange, Frederick (see also German Materialists) 14 (n 20), 113 (n 11)
Leiter, Brian 15, 16 (n 22), 25 (n 37), 39 (n 8), 224, 256, 259–63
Libet, Benjamin 123, 124 (n 21)
life 9–12, 18–19, 24–5, 27, 72–3, 75 (n 18), 78–81, 84, 89–103, 169–73, 175–7, 212–13, 219, 224–32, 259–63;
affirmation of 8–10, 81–5, 83 (n 44), 97–8, 169, 224–6, 236, 241–3;
see also affirmation, love
love 242;
of self 81–2, 197;
of fate 213;
see also amor fati
Lovibond, Sabina 206

masters 6, 16, 37, 66, 85, 136, 143–6, 151 (n 1), 168 (n 27)
May, Simon 148 (n 42)
meaning 24–28, 89, 96–102, 171–2, 227–9
Mele, Alfred 124 (n 21)
memory 37, 119 (n 16), 139–40, 182, 203
metaphysics 4 (n 4), 12–14, 44–5, 62–3, 66, 155 (n 7), 172 (n 32), 175–7, 249
Montaigne, Michel de 73–5, 86
morality 24, 37, 66, 163–5, 171, 197–8, 202–10, 260;
of custom 60–1, 128, 128 (n 4), 181–2, 193–4, 219–20;
critique of 53–4, 90, 131, 148, 223, 260;
decline of 101–3, 145;
errors promoted by 121, 228;
harmful to life 81, 104–6, 144;
naturalistic account of 116–21, 239, 247;
slave 37, 70, 89, 136, 205;
(*see also* masters; slave)
moral law 194, 198–202
Musil, Robert 153 (n 5)

Nagel, Thomas 115
naturalism 1, 14–29, 38, 41, 44–6, 58, 71–2 (n 9), 74 (n 36), 129–33, 141, 219, 223–6, 247–51, 259–60, 265 (n 7), 267
Nehamas, Alexander 154 (n 6), 224, 226, 236 (n 22)
nihilism 85, 89, 98–106

Owen, David 186 (n 4)

Pascal, Blaise 73–4, 86, 175
pessimism *see* nihilism
Pereboom, Derk 107
personhood 44–5, 127, 155, 169, 177, 203
perspectivism *see also* knowledge; truth
Plato 4, 5, 12, 89, 90, 98–100 (and n 26), 104, 106, 267
power 4–7, 9–10, 13–14, 18, 52, 54–62, 77, 84, 89–94, 110–11, 128–39, 142–6, 212
priest 26, 91, 105, 239
promise-making 36–7, 61–2, 71 (n 6), 89–94, 104, 119 (n 16), 138–42, 182–92, 194, 203–11
psychology 7, 13, 15–16, 26–8, 71–8, 107–8, 122–4, 132–3, 159–60, 173–4 (and n 36), 223, 224 (n 2), 239, 263–4
psychoanalysis 13, 14, 225 (n 7), 230–3
punishment 33–41, 47, 130–1, 140, 182

rank ordering 92, 228, 267 (n 8)
reactive 141, 151 (n 1)
  see also ressentiment
realism 1, 4, 6–7, 12, 159–61;
  about value
redemption 61, 97, 98, 102, 104
Reé, Paul 63
Reginster, Bernard 13 (n 18), 100 (n 26)
responsibility 34–44, 90, 104–5, 108, 111, 113–15, 119 (and n 16), 127–9, 146–9, 205–7, 216, 254 (n 4), 260 (n 6)
ressentiment 4, 26, 64, 91, 93, 96–7, 116, 165–9, 223, 239–41
repression 144, 231–5;
  see also Freud
'revaluation of values' 14, 57, 64–5, 92, 103, 131–2, 147–8
revenge 72 (n 8), 93, 97, 99, 104, 169, 240;
  see also ressentiment; Self
Richardson, John 38 (n 7), 224, 233 (n 18)
Ridley, Aaron 60, 108 (n 3), 113 (n 11), 206–7, 209, 215–17

Schelling F.W.J. *see* German Romantics
Schopenhauer, Arthur 10, 14, 27, 35–6, 51–2, 58 (n 17), 63, 69, 74, 249–50
science 7, 17–18, 21–3, 54, 75 (n 18), 123, 173–6, 223, 229–30, 249
Self 8, 2 (n 1), 2–15, 125, 132;
  creation/overcoming 8, 42, 78–9, 152, 156–7, 167, 214;
  unified 12–13, 16–21, 38, 45, 57–60, 72, 77, 79, 134–5, 169;
  knowledge 24, 61–5, 156, 193, 210, 230–5, 235 (n 21), 241 (and n 30), 243;
  see also 'giving style to one's character'
slave 53, 66, 132 (n 13), 141, 151 (n 1), 155–6, 168 (n 27), 239;
  see also master; morality
soul 2 (n 1), 3 (n 2), 5, 12–14, 44–5, 55, 59, 71, 75–7, 90, 98 (n 24), 101, 144, 176, 199–201, 248–9, 253–9, 263–4, 267
sovereign individual 8, 21, 36–46, 60–4, 80 (n 35), 96, 104, 119 (n 16), 128–9, 142–3, 151–2, 156, 168, 181–94, 203–9, 211–212, 216–17;
  and freedom 36–46, 60–4, 128–9, 142–3
Stirner, Max 10–13
Strawson, Galen 107, 115, 161 (n 17)
suffering 4, 18, 24–8, 83, 89–90, 97–105, 152, 168 (n 27)
Sykes, Chris 45 (n 14)

Transcendentalism 1, 5, 9–14, 19–23, 28–9, 57, 152, 154–5, 169, 171 (n 30)
truth 22 (n 33), 94 (n 17), 94–6, 99–101;
  value of 71, 81, 102–3, 165, 169–70, 173–7, 216–18, 230, 241 (n 30), 265 (n 7);
  see also perspectivism; will to truth
types 20, 24, 61, 75, 146, 224;
  facts 113 (n 11), 115–23, 125, 223;
  doctrine of 115–22

unconscious (non-conscious) 13, 53, 92, 115–8, 118 (n 15), 122, 124–5, 138, 140, 158 (n 13), 231–5
unified self (see self, unified)

value 71–4;
  creation 23–6, 62, 64–6, 89–92;
  freedom and 69, 129–31;

value (*cont.*)
life and 98–103;
Nietzsche's conception of 8–10, 15–17, 61, 71 (n 6), 131;
see also 'revaluation of values'
valuing 8, 15, 59, 75 (n 18), 136, 171
Velleman, J. David 72 (n 10)

Wagner, Richard 34–5, 42–3. 82 (and n 42), 94–96, 135 (n 20)
Watson, Gary 34 (n 3), 157 (n 11 and 12), 161
Wegner, Daniel M. 107, 122–24, 255
Will 43, 91 (and n 9), 97–9, 103–4, 202–3, 247, 254–60, 263–6;
empirical psychology and 107–8, 122–4;
as epiphenomenal 247–8, 254–9, 260 (n 6);
phenomenology of 108–16, 118, 121–4, 247–54, 260–3;
see freedom; free will
Will to power 4, 9, 17, 20, 74, 84–6, 93, 129, 267
Will to truth; 14–17, 144, 170–7, 229–30;
see also truth, value of
Williams, Bernard 72 (n 7), 208
Wittgenstein, Ludwig 167